Ano 2023

INFERÊNCIA NO ENSINO MÉDIO:

UMA INTRODUÇÃO AOS TESTES DE HIPÓTESE

PAULO ROBERTO CONSTANTINO JUNIOR

Ano 2023

INFERÊNCIA NO ENSINO MÉDIO:

UMA INTRODUÇÃO AOS TESTES DE HIPÓTESE

PAULO ROBERTO CONSTANTINO JUNIOR

CIÊNCIAS EXATAS E DA TERRA E ENGENHARIAS

Inferência no ensino médio: uma introdução aos testes de hipótese

Diagramação: Camila Alves de Cremo
Correção: Maiara Ferreira
Indexação: Amanda Kelly da Costa Veiga
Revisão: O autor
Autor: Paulo Roberto Constantino Junior

Dados Internacionais de Catalogação na Publicação (CIP)

C758 Constantino Junior, Paulo Roberto
Inferência no ensino médio: uma introdução aos testes de hipótese / Paulo Roberto Constantino Junior. – Ponta Grossa - PR: Atena, 2023.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-65-258-0793-5
DOI: https://doi.org/10.22533/at.ed.935222312

1. Inferência (Lógica). 2. Estatística. 3. Probabilidade. I. Constantino Junior, Paulo Roberto. II. Título.

CDD 401.4

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

Atena Editora
Ponta Grossa – Paraná – Brasil
Telefone: +55 (42) 3323-5493
www.atenaeditora.com.br
contato@atenaeditora.com.br

DECLARAÇÃO DO AUTOR

O autor desta obra: 1. Atesta não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao conteúdo publicado; 2. Declara que participou ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certifica que o texto publicado está completamente isento de dados e/ou resultados fraudulentos; 4. Confirma a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhece ter informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autoriza a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.

CIÊNCIAS EXATAS E DA TERRA E ENGENHARIAS

DECLARAÇÃO DA EDITORA

Dissertação de mestrado apresentada ao Centro de Matemática, Computação e Cognição para obtenção do título de Mestre

ESTE EXEMPLAR CORRESPONDE A VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO PAULO ROBERTO CONSTANTINO JUNIOR, E ORIENTADA PELO PROF. DR. ANDRÉ RICARDO OLIVEIRA DA FONSECA.

Dedico este trabalho ao meu filho Murilo.

AGRADECIMENTOS

Agradeço em primeiro lugar a minha mãe. Ela sempre esteve ao meu lado em todos os momentos, me inspirando e apoiando. Foi sempre meu maior exemplo de vida. Agradeço por toda a compreensão, por entender as ausências que muitas vezes de fizeram necessárias.

A minha querida esposa Áurea, pelo apoio, compreensão, pelos sonhos compartilha- dos. Por entender que nem sempre podia estar presente em certos momentos.

Ao meu filho Murilo e a Brenda, minha filha do coração, por estarem sempre ao meu lado, apoiando esse processo de construção desse projeto de vida.

Ao meu pai que, apesar da distância esteve sempre presente em pensamentos.

Aos meus irmãos Daniela e Carlos por sempre me apoiarem, estarem dispostos a ajudar no que fosse preciso e entender os momentos em que tive que me fazer ausente para estudar e me dedicar a esse sonho.

Aos meus familiares e amigos, agradeço por todo apoio. Em especial a minha prima Débora, pelo auxílio nas traduções.

Aos colegas da Turma do PROFMAT, em especial ao meu amigo Rodolfo, pela com- preensão, por compartilhar os momentos de dificuldade, as alegrias com os sucessos obtidos, a ajuda mútua, as caronas, os momentos compartilhados de estudo, meu muito obrigado.

A Direção, Coordenação, corpo Docente e funcionários da E. E. Professora Therezi- nha Closa Eleutério, em especial aos alunos do 3o ano A.

A Professora Sueli pela revisão ortográfica, e a Professora Elaine que gentilmente me cedeu alguns horários de aula para realização das atividades.

A UFABC, por ter me dado a oportunidade de crescimento profissional e aberto as portas de sua instituição para concretização e realização desse sonho.

Ao corpo Docente da UFABC pelas aulas ministradas, que me permitiram uma evo- lução pessoal e profissional.

Ao meu orientador Professor Doutor André Ricardo Oliveira da Fonseca, que sempre me orientou de forma positiva, com muita disposição e presteza nos retornos do meu Projeto.

A CAPES pelo apoio financeiro, sem o qual não seria possível a dedicação e conclusão deste trabalho.

“Quando você elimina o impossível, o que sobra por mais incrível que pareça só pode ser a verdade.”

(Sir Arthur Conan Doyle)

SUMÁRIO

SUMÁRIO

SUMÁRIO

RESUMO

No mundo contemporâneo é comum constantes pesquisas em diversos âmbitos, tanto sociais, quanto econômicos, entre outros. Para tais pesquisas é fundamental a coleta de dados, organizar os dados, como também construir tabelas e gráficos es- tatísticos, entretanto é inadmissível não haver uma interpretação consistente sobre os resultados. Desta forma, o objetivo primordial deste trabalho é apresentar aos alunos do Ensino Médio, especificamente os do 3o ano, a teoria da inferência estatística e aplicações, para que eles possam, num nível elementar, compreender os meios de ob- tenção de uma amostra e as conclusões possíveis sobre a respectiva população. Assim, estimulando os educandos em buscar constantemente informações sobre pesquisas es- tatísticas, as quais estarão presentes em vários momentos da sua vida em sociedade.
PALAVRAS-CHAVE: Curva normal, inferência estatística, probabilidade.

ABSTRACT

In the contemporary world it is common to come across frequent research from various scopes, both social and economical amongst others. For such research, it is vital to collect data, organize it as well as put together statistical charts and graphs. However it is unacceptable that there is no consistent interpretation about the results. Hence, the primary focus of this work is present to High School students, specifically the seniors, the theory of statistical inference and its applications, so that they can, on an elementary level, understand the means of obtaining a sample and the possible conclusions about a respective population. Thus, we expect to stimulate the students to constantly seek information about statistical research, which will be present in many different moments in their lives as part of society.
KEYWORDS: Normal curve, statistical inference, probability.

INTRODUÇÃO

A estatística está cada vez mais presente em nosso cotidiano através dos meios de comunicação em massa, nas indústrias e no comércio. Na escola diferentes disciplinas utilizam gráficos e tabelas de dados estatísticos, além da própria matemática. A geografia faz uso da estatística nos estudos comparativos de dados populacionais, a biologia na análise dados coletados, a história e a sociologia no tratamento da informação.

O presente trabalho justifica-se pela crescente necessidade de produtos com qualidade em um mercado atual e dinâmico, como também maior agilidade na entrega do produto final ao consumidor, que cada vez está mais exigente.

Assim, dentre vários objetivos da educação básica no Brasil, sem dúvida o de preparar os jovens para este mercado atual é imprescindível.

No primeiro capítulo é apresentado o desenvolvimento da estatística como área de conhecimento, definindo as principais medidas de posição e de dispersão da estatística descritiva; analisando as diferentes distribuições de probabilidade, algo de grande importância para a estatística inferencial.

As definições e os teoremas, como o Teorema do Limite Central, com suas demonstra- ções, no ramo da inferência estatística encontram-se no segundo capítulo. Incluem-se, ainda, os intervalos de confiança e os testes de hipótese para a média, a variância e proporções.

O terceiro capítulo destina-se a atividades propostas aos alunos, dentre as quais: a aproximação da distribuição binomial à normal e a inferência de uma amostra obtida pelos próprios alunos. Assim, orientando aos educandos do 3o ano do Ensino Médio para que sejam protagonistas do próprio conhecimento.

A aplicação das atividades em uma sala de aula do 3o ano do Ensino Médio da escola estadual Professora Therezinha Closa Eleutério, na cidade de Guarulhos, em São Paulo, encontra-se no quarto capítulo, no qual se registra o diálogo entre professor e aluno.

FUNDAMENTAÇÃO TEÓRICA

1 I SÍNTESE HISTÓRICA

Muitos foram os nomes ao longo da história que contribuíram para a Teoria Estatística, desde Gauss e Legendre com a descoberta do método dos mínimos quadrados, até Yule, que o relacionou à regressão linear.

Primeiramente com o surgimento das primeiras cidades, houve a necessidade de se contar a quantidade de habitantes, de inimigos, dos animais nos rebanhos, inclusive da produção agrícola; para isso utilizaram o censo, o qual também foi utilizado pelos povos da antiguidade, justamente pela necessidade da distribuição de terras para o plantio, da manutenção do exército, da coleta de impostos, entre outros.

John Graunt que no ano de 1662, publicou estudos estatísticos sobre os índices de mortalidade ocorridos em Londres, dando ênfase aos dados obtidos; bem como realizou os primeiros trabalhos no desenvolvimento da teoria das probabilidades, trabalhos primeiramente iniciados com a troca de correspondências entre os matemáticos Pascal e Fermat. Vale ressaltar que as primeiras pesquisas de Graunt se restringiam sobretudo ao seu interesse inicial, ou seja, os jogos de azar.

Jacob Bernoulli (1654 – 1705), que também contribuiu à probabilidade, mostrou através de seus estudos que quanto maior o número de observações, menor a incerteza no resultado.

Abraham De Moivre (1667 – 1754) publicou, em 1733, o que hoje é conhecido como aproximação normal à binomial. Também foi o primeiro a trabalhar com a curva característica em forma de sino da distribuição normal, a curva de frequência normal

$$y = ce^{-hx^2},$$

com c e h constantes.

Thomas Simpson (1710 – 1761) em seus estudos sobre o raciocínio reverso das proba- bilidades envolvendo a distribuição binomial, voltou seu foco aos erros de observação, não nas observações.

Thomas Bayes (1701 – 1761) e Pierre-Simon de Laplace (1749 – 1827) deram de forma independente, uma prova formal do que hoje é conhecido como teorema de Bayes, demonstrado anteriormente, mas não formalmente, por Jacob Bernoulli.

Florence Nightingale (1820 – 1910) , enfermeira durante a Guerra da Criméia (1854 – 1856) mostrou, por meio de gráficos, que as péssimas condições de higiene dos hospitais na época matavam mais soldados do que a própria guerra.

Problemas envolvendo as órbitas de Júpiter e Saturno e a órbita da Lua, ganharam grande destaque no século *XVIII*. Tais problemas levaram à introdução e ao desenvolvimento de métodos de resolução das equações, resultantes das observações destas

órbitas.

Em 1749, Leonhard Euler (1707 – 1783) publicou seus estudos sobre as órbitas de Júpiter e Saturno e Johann Tobias Mayer (1723 – 1762) publicou, no ano seguinte, seus estudos sobre a órbita da Lua.

Euler trabalhou como matemático, com base em observações realizadas anteriormente, enquanto que Mayer realizou diversas observações, de modo que obteve su- cesso em encontrar uma solução estatística para o problema da órbita da Lua; porém Euler não.

O método de Mayer levava a resultados aceitáveis e, mesmo não sendo o melhor método para se encontrar as soluções das equações lineares, influenciou os estudos seguintes em estatística.

Além das órbitas de Júpiter e Saturno e da órbita da Lua, outro grande problema do século *XVIII*, estudado por Boscovich, era determinar o formato da Terra. Antes de Boscovich outros já haviam dividido a Terra em arcos e sabiam que ela era achatada no pólos, o problema era saber quanto.

Por isso Boscovich comparou a medida de arcos bem separados em latitude, pois em medidas muito próximas os erros poderiam ser maximizados, como resultado apresen- tou uma descrição geométrica de seu método, enquanto Laplace, em 1789, fez uma descrição algébrica do método de Boscovich.

Adrien-Marie Legendre (1752 – 1833) propôs, em 1805, em seu trabalho sobre a órbita dos cometas, o método dos mínimos quadrados, no qual os erros resultantes poderiam ser minimizados tomando-se a soma dos quadrados dos desvios.

Legendre, diante dos problemas das órbitas e da medição da Terra, na mesma época em que os métodos foram desenvolvidos, percebeu que o método dos mínimos quadrados não se resumia à órbita dos cometas; então Laplace formalizou uma prova da teoria dos mínimos quadrados.

Carl Friedrich Gauss (1777 – 1855) investigou os erros de observação das órbitas dos planetas em termos probabilísticos, dando origem ao modelo de distribuição normal, conhecido como distribuição gaussiana. Da mesma maneira que a curva normal é conhecida por curva de Gauss.

Gauss descobriu o método dos mínimos quadrados em 1795, não o publicando antes de 1805, criando uma disputa pela prioridade da descoberta, entretanto, não há indícios de que Gauss viu seu potencial antes dos trabalhos de Legendre serem publicados.

O método dos mínimos quadrados se mostrou útil nos estudos astronômicos e geodésicos, no entanto não havia ainda aplicação às ciências sociais. [18]

Adolphe Quetelet (1796 – 1874) estudou vários aspectos sociais, desta forma encon- trou um padrão nas distribuições, sendo que os dados seguiam a curva normal. Foi falho, porém contribuiu em avaliar e encontrar meios de classificar dados para análise, fazendo-o apenas em situações específicas.

Siméon-Denis Poisson (1781 – 1840) publicou, em 1837, um livro sobre a aplicação

das probabilidades em vereditos de um júri, o qual introduziu um modelo de distribui- ção de probabilidades, a distribuição de Poisson.

Wilhelm Lexis (1837 – 1914) publicou, no ano de 1876, séries estatísticas sobre taxas de natalidade e mortalidade. Três anos depois, publicou um artigo sobre a estabilidade de séries estatísticas, encontrando séries estáveis determinadas mais geneticamente do que socialmente.

Gustav Theodore Fechner (1801 – 1887) criou uma metodologia quantitativa em psicologia, seu método consistia em estímulos e respostas, logo após diversas repetições, contava-se a quantidade de respostas certas e erradas.

Hermann Ebbinghaus (1850 – 1909) utilizando as ideias de Fechner, desenvolveu um argumento para justificar o tratamento dos dados, baseado na distribuição normal.

Francis Galton (1822 – 1911) aplicou a estatística aos estudos sobre hereditariedade, para mostrar que as variações tendem a retornar à população geral, entretanto não conseguiu desenvolver matematicamente suas próprias ideias.

Francis Ysidro Edgeworth (1845 – 1926) traduziu as ideias de Galton para a forma matemática, no estudo das estatísticas sociais e econônicas, aproximando a distribui- ção da população à distribuição normal.

Karl Pearson (1857 – 1936) usou a correlação nas estatísticas sociais, encontrando diversidade onde Galton encontrou uniformidade, limitando-se exclusivamente aos estudos das ciências sociais.

George Udny Yule (1871 – 1951) não limitou o estudo da estatística à sua área de conhecimento, a biologia; relacionando a regressão linear aos mínimos quadrados.

O geneticista britânico Sir Ronald Aylmer Fischer (1890 – 1962) mostrou a importância do planejamento experimental e desenvolveu a análise da variância, ANOVA.

2 | ESTATÍSTICA DESCRITIVA

A coleta, a organização, redução e representação dos dados, constituem o ramo da estatística descritiva, que tem por interesse os próprios dados, não preocupando-se em tirar conclusões.

2.1 Variáveis

O conhecimento dos tipos de variáveis envolvidas determina o tipo de tratamento que será dado às informações obtidas. Essas variáveis podem ser classificadas em qualitativas ou quantitativas.

As variáveis qualitativas, tais como rótulos, qualidades ou entradas não numéricas, podem ser nominais, quando os dados a serem considerados não podem ser ordenados; ou ordinais, em que os dados podem ser ordenados.

As variáveis quantitativas podem ser discretas, sendo que os dados, obtidos geralmente através de uma contagem, formam um conjunto enumerável; ou contínuas, na qual

os dados pertencem a um intervalo de números reais, resultando de medições.

2.2 Medidas de Posição

As medidas de posição são valores representativos de toda a série de dados. Uma medida de posição central descreve uma entrada central do conjunto de dados, sendo as mais usuais: a média, a mediana e a moda.

Definição 1.1. A média aritmética da variável X é a soma das entradas x_1, x_2, . . . , x_n dividida pelo número n de entradas.

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Definição 1.2. A mediana da variável X é o valor que está no centro do conjunto ordenado de n entradas, x_1, x_2, . . . , x_n, dividindo-o em duas partes iguais.

$$md(X) = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ ímpar;} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ par.} \end{cases}$$

Definição 1.3. A moda da variável X é a entrada mais frequente de um conjunto de dados.

2.3 Medidas de Dispersão

A dispersão mede o quanto um conjunto de dados está afastado de sua média; se os valores estão próximos, ela é pequena e se estão espalhados, ela é grande.

Dada uma variável X, de entradas x_1, x_2, . . . , x_n e média $\bar{x}$, os desvios das entradas em relação à média são as diferenças:

$$x_1 - \overline{x}, x_2 - \overline{x}, \ldots, x_n - \overline{x}$$

A soma destes desvios é igual a zero.

$$\begin{aligned} \sum_{i=1}^{n}(x_i - \overline{x}) &= (x_1 - \overline{x}) + (x_2 - \overline{x}) + \cdots + (x_n - \overline{x}) \\ &= (x_1 + x_2 + \cdots + x_n) - n \cdot \overline{x} \qquad (1.1) \\ &= n \cdot \overline{x} - n \cdot \overline{x} \qquad (1.2) \\ &= 0. \end{aligned}$$

A passagem de 1.1 para 1.2 segue da Definição 1.1.

Para que a soma dos desvios não seja igual a zero, devemos considerar os valores absolutos dos desvios ou o quadrado dos desvios. Desta forma,

$$\sum_{i=1}^{n}|x_i - \overline{x}|$$

ou

$$\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Assim, definimos o desvio médio, a variância e o desvio padrão, de uma variável *X*.

Definição 1.4. O desvio médio é a média da soma dos valores absolutos dos desvios.

$$dm(X) = \frac{\sum_{i=1}^{n}|x_i - \overline{x}|}{n}.$$

Definição 1.5. A variância é a média da soma dos quadrados dos desvios.

$$var(X) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}.$$

Definição 1.6. O desvio padrão é a raiz quadrada da variância.

$$dp(X) = \sqrt{var(X)}.$$

3 | PROBABILIDADE

3.1 Espaço amostral e evento

Definição 1.7. Espaço amostral é o conjunto de todos os resultados possíveis de um experimento, sendo representado pela letra *S*. Evento é um subconjunto do espaço amostral.

Considerando dois eventos *E* e *F* quaisquer de um espaço amostral *S*, temos as seguintes operações com conjuntos:

- A união é o conjunto dos elementos que pertencem a *E* ou a *F*, ou a ambos:

$$E \cup F = \{x : x \in E \text{ ou } x \in F\};$$

- A interseção é o conjunto dos elementos que pertencem a *E* e a *F* simultaneamente:

$$E \cap F = \{x : x \in E \text{ e } x \in F\};$$

- O complemento de *E*, denotado por E^C, é o evento formado por todos os elemen- tos do espaço amostral que não estão contidos em *E*:

$$E^C = \{x : x \notin E\}.$$

3.2 Definição e algumas probabilidades importantes

A cada evento *E* do espaço amostral *S* está associado um número, denotado por *P*(*E*), chamado de probabilidade do evento *E*.

Definição 1.8. Uma probabilidade é uma função P com domínio em S, que associa a cada evento E um número $P(E)$ que satisfaz as seguintes propriedades:

1. $0 \leq P(E) \leq 1$, para todo evento E;

2. $P(S) = 1$;

3.Se os eventos E_1, E_1, . . ., são mutuamente excludentes dois a dois, isto é, eventos para os quais $E_i \cap E_j = \emptyset$ quando $i \neq j$, então

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i). \tag{1.3}$$

As propriedades apresentadas na Definição 1.8 são conhecidas como Axiomas de Probabilidade. O terceiro axioma é conhecido como Axioma da Aditividade Contável.

Teorema 1.9. *Se P é uma função de probabilidade e E é um subconjunto de S, então:*

(i) $P(EC) = 1 - P(E)$*;*

(ii) $P(\emptyset) = 0$*, onde* $\emptyset$ *é o conjunto vazio.*

Demonstração. Inicialmente, provaremos (i).

Os conjuntos E e E^C são tais que $S = E \cup E^C$. Portanto, pelo Segundo Axioma das Probabilidades,

$$P(E \cup E^C) = P(S) = 1 \tag{1.4}$$

Como, E e E^C são mutuamente excludentes, temos, pelo Terceiro Axioma das Probabilidades,

$$P(E \cup E^C) = P(E) + P(E^C) \tag{1.5}$$

De (1.4) e (1.5), resulta (a).

Para provar (ii), utilizamos o fato de que $S = S \cup \emptyset$, e que S e $\emptyset$ são mutuamente excludentes, então, por um raciocínio análogo, temos,

$$1 = P(S) = P(S \cup \emptyset) = P(S) + P(\emptyset),$$

e, portanto, $P(\emptyset) = 0$.

3.3 Contagem

A seguir serão apresentadas algumas técnicas de contagem, que serão úteis no cál- culo de probabilidades.

Teorema 1.10 (Teorema Fundamental da Contagem)**.** *Se uma escolha consiste em k passos, a i-ésima pode ser realizada de* n_i *modos,* $i = 1, \ldots, k$*, então a escolha total pode ser feita de* $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ *maneiras.*

Demonstração. A demonstração será feita pelo Princípio da Indução Finita.

A primeira escolha pode ser feita de n_1 modos, e para cada um desses modos, temos n_2 opções para a segunda escolha. Desse modo, podemos realizar as escolhas de

$$\underbrace{(1 \cdot n_2) + (1 \cdot n_2) + \cdots + (1 \cdot n_2)}_{n_1 \text{ termos}} = n_1 \cdot n_2$$

modos, provando o teorema para $k = 2$.

Supondo que o teorema seja válido para k escolhas, então, a quantidade de maneiras de se fazer a escolhas é:

$$n_1 \cdot n_2 \cdot \ldots \cdot n_k.$$

O objetivo agora, é mostrar que o teorema seja válido para n_{k+1} escolhas. Assim,

$$\underbrace{(n_1 \cdot n_2 \cdot \cdots \cdot n_k \cdot 1) + \cdots + (n_1 \cdot n_2 \cdot \cdots \cdot n_k \cdot 1)}_{n_{k+1} \text{ termos}} = n_1 \cdot n_2 \cdot \cdots \cdot n_k \cdot n_{k+1}.$$

Portanto, o teorema é válido para k escolhas.

A seguir, será apresentada a definição de fatorial de um número inteiro positivo.

Definição 1.11. Para um número inteiro positivo n, definimos $n!$ como o produto de todos os números inteiros positivos menores que ou iguais a n.

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1.$$

Definimos, ainda, $0! = 1$.

Definição 1.12. Para números inteiros não negativos, n e r, com $r \leq s$, definimos $\binom{n}{r}$, como

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}$$

3.4 Tipos de Probabilidades

Supondo que todos os resultados de um experimento, no qual o espaço amostral $S = \{s_1, s_2, \ldots, s_N\}$ é um conjunto finito, são igualmente prováveis. Podemos dizer que

$$P(\{s_1\}) = P(\{s_2\}) = \cdots = P(\{s_N\}).$$

Então,

$$P(\{s_i\}) = \frac{1}{N}, \text{ com } i = 1, \ldots, N.$$

Do Terceiro Axioma das Probabilidades, resulta que, para cada evento E,

$$P(E) = \frac{n(E)}{n(S)} \tag{1.6}$$

onde $n(E)$ e $n(S)$ são, respectivamente, os números de elementos em E e S.

A equação 1.6 é conhecida como probabilidade clássica.

Um outro tipo de probabilidade, conhecida como probabilidade empírica, é baseada na frequencia com que um evento ocorre a longo prazo. Considerando um evento E, nestas condições, temos,

$$P(E) = \frac{\text{Frequencia do evento } E}{\text{Frequencia total}}$$

3.5 Probabilidade condicional e independência

A probabilidade de um evento ocorrer, dado que outro evento já ocorreu, é denominada probabilidade condicional.

Definição 1.13. Dados dois eventos E e F, $P(F) > 0$, definimos a probabilidade condi- cional de E dado F, representada por $P(E|F)$, como

$$P(E|F) = \frac{P(E \cap F)}{P(F)}. \tag{1.7}$$

A relação a seguir, obtida de (1.7), é a chamada regra do produto das probabilidade,

$$P(E \cap F) = P(F)P(E|F).$$

Podemos também escrever,

$$P(F \cap E) = P(E)P(F|E),$$

devido à simetria de (1.7).

Como $P(E \cap F) = P(F \cap E)$, temos

$$P(F)P(E|F) = P(E)P(F|E).$$

Ou seja,

$$P(E|F) = \frac{P(E)P(F|E)}{P(F)}. \tag{1.8}$$

Que é a versão mais simples da Regra de Bayes.

Considere os eventos E e F, o evento F pode ser escrito como

$$F = (F \cap E)(F \cap E^C).$$

Como $(F \cap E)$ e $(F \cap E^C)$ são mutuamente excludentes, temos da Equação (1.3), que

$$P(F) = P(F \square E) + P(F \square E^C)/nonumber \tag{1.9}$$

$$= P(E)P(F|E) + P(E^C)P(F|E^C). \tag{1.10}$$

Substituindo (1.10) em (1.8), obtemos

$$P(E|F) = \frac{P(E)P(F|E)}{P(E)P(F|E) + P(E^C)P(F|E^C)} \tag{1.11}$$

Podemos generalizar (1.11).

Sejam E_1, E_2, . . . uma partição de espaço amostral, e F um conjunto qualquer, então a probabilidade de ocorrência de E_i, supondo-se a ocorrência de F, é dada por

$$P(E_i|F) = \frac{P(E_i)P(F|E_i)}{\sum_{j=1}^{\infty} P(E_j)P(F|E_j)}, \tag{1.12}$$

para todo $i = 1, 2, \ldots$.

Pode ocorrer de um evento F não ter influência sobre um evento E. Neste caso, temos

$$P(E|F) = P(E).$$

Então, pelo Regra de Bayes, obtemos, dessa forma

$$P(F|E) = \frac{P(F)P(E|F)}{P(E)} = \frac{P(F)P(E)}{P(E)} = P(F).$$

Como $P(E \cap F) = P(E)P(F|E)$, temos

$$P(E \cap F) = P(E)P(F).$$

Motivando, assim, a seguinte definição.

Definição 1.14. Dois eventos E e F são estatisticamente independentes se

$$P(E \cap F) = P(E)P(F). \tag{1.13}$$

4 | VARIÁVEIS ALEATÓRIAS

Definição 1.15. Uma variável aleatória é uma função X definida no espaço amostral S que assume valores no conjunto dos números reais.

Existem dois tipos de variáveis aleatórias:

- Variáveis aleatórias discretas, quando a função X assume valores em um conjunto enumerárel de pontos da reta;
- Variáveis aleatórias contínuas, quando a função X assume valores em um intervalo no conjunto dos números reais.

4.1 Função de distribuição acumulada

A função de distribuição acumulada (f.d.a.) de X é uma função associada a cada uma das variáveis aleatórias X.

Definição 1.16. Dada uma variável aleatória X, a função acumulada (f.d.a.), denotada por $F_X(x)$, é dada por

$$F_X(x) = P(X \leq x), \text{ para todo } x.$$

4.2 Função de probabilidade ou densidade de probabilidade

A função de probabilidade (f.p.), no caso da variável aleatória discreta, ou função densidade de probabilidade (f.d.p.), no caso da variável aleatória contínua, é uma função associada a uma variável aleatória X e a sua (f.d.a.).

Definição 1.17. A função de probabilidade (f.p.), de uma variável aleatória discreta X, é dada por:

$$f_X(x) = P(X = x), \text{ para todo } x.$$

Definição 1.18. A função densidade de probabilidade (f.d.p.), de uma variável aleatória contínua X, é a função que satisfaz:

$$F_X(x) = \int_{-\infty}^{x} f(t)dt, \text{ para todo } x.$$

A área compreendida entre dois valores a e b quaisquer, sob a curva de $f_X(x)$, nos fornece a probabilidade de uma variável aleatória com distribuição contínua. Podemos escrever

$$P(a \leqslant X \leqslant b) = \int_{a}^{b} f_X(x)dx = F_X(b) - F_X(a). \tag{1.14}$$

4.3 Valor esperado

A média, ou valor esperado, de uma variável aleatória é um número que representa um valor que se espera de uma observação da variável aleatória.

Definição 1.19. O valor esperado de uma variável aleatória X, de uma função $g(X)$, representado por $Eg(X)$, é

$$Eg(X) = \begin{cases} \sum g(X)P(X = x), & \text{se } X \text{ for uma variável discreta;} \\ \int_{-\infty}^{\infty} g(X)f_X(x)dx, & \text{se } X \text{ for uma variável contínua.} \end{cases}$$

4.4 Momentos

Definição 1.20. Para cada n inteiro, o n-ésimo momento de X é definido como

$$\mu' n = EX^n.$$

O n-ésimo momento central de X é dado por

$$\mu_n = E(X - \mu)^n,$$

onde $\mu = \mu_1' = EX$.

4.5 Variância e desvio padrão

A variância é o segundo momento central, podemos definir a variância e o desvio padrão de uma variável aleatória.

Definição 1.21. A variância de uma variável aleatória X, com média μ, representada por $Var(X)$, é

$$Var(X) = E(X - EX)^2. \tag{1.15}$$

O desvio padrão de uma variável aleatória X é a raiz quadrada positiva de $Var(X)$, representado por $DP(X)$.

$$DP(X) = \sqrt{Var(X)}.$$

O desvio padrão mede a dispersão da população em torno da média.

Podemos simplificar o cálculo a variância por meio da fórmula deduzida a seguir:

$$Var(X) = E(X - EX)^2$$
$$= E[(X^2 - 2XEX + (EX)^2] \tag{1.16}$$
$$= E(X^2) - 2(EX)^2 + (EX)^2 \tag{1.17}$$
$$=E(X^2) - (EX)^2 \tag{1.18}$$

Usando o fato de EX ser uma constante, podemos escrever

$$E(XEX) = (EX)(EX) = (EX)^2$$

quando passamos de (1.16) para (1.17).

5 I DISTRIBUIÇÕES DE PROBABILIDADE DISCRETAS

5.1 Distribuição uniforme discreta

Nesta distribuição cada um dos resultados 1, . . . , N, tem a mesma probabilidade de ocorrer. Uma variável aleatória discreta X tem distribuição uniforme se

$$P(X = x\ N) = \frac{1}{N}, \quad x = 1, \ldots, N.$$

Para o cálculo da média e da variância de X, podemos fazer uso dos seguintes teore- mas:

Teorema 1.22. *A soma dos n primeiros números naturais é* $\frac{n(n+1)}{2}$.

Demonstração. A demonstração será feita pelo Princípio da Indução Finita.

Seja $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

Para $n = 1$, a igualdade é verdadeira, pois $\frac{1(1+1)}{2} = 1$. Suponha que ela seja verdadeira, para algum $n \in \mathbb{N}$. Então,

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Adicionando $(n + 1)$, a ambos os lados, obtemos:

$$\sum_{i=1}^{n} i + (n+1) = \frac{n(n+1)}{2} + (n+1).$$

Resolvendo,

$$\sum_{i=1}^{n+1} i = \frac{n(n+1)}{2} + \frac{2(n+1)}{2} = \frac{(n+1)(n+2)}{2}.$$

Portanto,

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

é verdadeira para todo $n \in \mathbb{N}$.

Teorema 1.23. *A soma dos quadrados dos n primeiros números naturais é* $\frac{n(n+1)(2n+1)}{6}$.

Demonstração. A demonstração será feita pelo Princípio da Indução Finita.

Seja $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

A igualdade é verdadeira para $n = 1$, pois $\frac{1(1+1)(2 \cdot 1+1)}{6} = 1 = 1^2$.

Suponha que ela seja verdadeira, para algum $n \in \mathbb{N}$. Então,

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

Adicionando $(n + 1)^2$, a ambos os lados, obtemos:

$$\sum_{i=1}^{n} i^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2.$$

Resolvendo, temos,

$$\sum_{i=1}^{n+1} i^2 = \frac{(n+1)[n(2n+1)+6(n+1)]}{6} = \frac{(n+1)(2n^2+7n+6)}{6} = \frac{(n+1)(n+2)(2n+3)}{6}.$$

Portanto,

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

é verdadeira para todo $n \in$ N.

Logo, o valor esperado é dado por:

$$EX = \sum_{x=1}^{N} xP(X = x) = \frac{1}{N}\sum_{x=1}^{N} x.$$

Do Teorema 1.22, temos

$$\frac{1}{N}\sum_{x=1}^{N} x = \frac{N+1}{2}$$

então,

$$EX = \frac{N+1}{2}. \tag{1.19}$$

Como,

$$EX^2 = \frac{1}{N}\sum_{x=1}^{N} x^2.$$

Do Teorema 1.23, obtemos,

$$\frac{1}{N}\sum_{x=1}^{N} x^2 = \frac{(N+1)(2N+1)}{6}$$

logo,

$$EX^2 = \frac{(N+1)(2N+1)}{6}. \tag{1.20}$$

Portanto, de (1.19) e (1.20), temos que a variância é

$$\begin{aligned} Var(X) &= EX^2 - (EX)^2 \\ &= \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 \\ &= \frac{(N-1)(N+1)}{12}. \end{aligned} \tag{1.21}$$

5.2 Distribuição binomial

Na distribuição binomial um experimento apresenta apenas dois resultados possíveis, podendo ser classificado como sucesso ou fracasso.

Definimos a seguir a variável aleatória de Bernoulli.

Definição 1.24. Uma variável aleatória X, que assume os valores 0, classificado como fracasso, e 1, classificado como sucesso, cuja função de probabilidade é

$$P(X = x) = p^x(1 - p)^{1-x}, \text{ com } x = 0, 1$$

onde p é a probabilidade de se obter sucesso, é chamada de variável aleatória de Bernoulli.

O valor esperado e a variância da variável aleatória de Bernoulli, são:

$$EX = 1p^1(1 - p)^{1-1} + 0p^0(1 - p)^{1-0}.$$

Logo

$$EX = p. \tag{1.22}$$

Da Equação (1.15), obtemos:

$$Var(X) = E(1 - p)^2 + E(0 - p)^2 = (1 - p)^2\, p + (0 - p)^2(1 - p).$$

Portanto

$$Var(X) = p(1 - p). \tag{1.23}$$

Suponhamos que sejam realizados n ensaios de Bernoulli, tal que n é previamente determinado, todos idênticos e independentes, com probabilidade p de sucesso e $1 - p$ de fracasso.

Se em n ensaios, obtemos x sucessos, então X é uma variável aleatória binomial com parâmetros n e p, ao qual denotamos $X \sim Binomial(n, p)$.

Sua função de probabilidade é dada por:

$$P(X = x) = \binom{n}{x} p^x(1 - p)^{n-x}, \text{ com } x = 0, \ldots, n. \tag{1.24}$$

Para calcularmos o valor esperado e a variância de uma variável aleatória de binomial, precisamos do Teorema binomial.

Antes, porém, se faz necessário o uso de uma igualdade muito importante.

Teorema 1.25. *Para números inteiros não negativos, n e r, com* $1 \leq r \leq s$*, temos:*

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \tag{1.25}$$

Demonstração. Da Definição 1.12, podemos escrever a igualdade da seguinte maneira:

$$\begin{aligned}
\binom{n-1}{r-1} + \binom{n-1}{r} &= \frac{(n-1)!}{(n-r)!\,(r-1)!} + \frac{(n-1)!}{(n-r-1)!\,r!} \\
&= \frac{(n-1)!}{(n-r)(n-r-1)!\,(r-1)!} + \frac{(n-1)!}{(n-r-1)!\,r(r-1)!} \\
&= \frac{n(n-1)! + (n-r)(n-1)!}{(n-r)(n-r-1)!\,r(r-1)!} \\
&= \frac{n!}{(n-r)!\,r!} = \binom{n}{r}.
\end{aligned}$$

Teorema 1.26 (Teorema binomial)**.** *Para números reais x e y e inteiros não negativos n, temos:*

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}. \tag{1.26}$$

Demonstração. A demonstração foi extraída de [15].

A demonstração será pelo Princípio da Indução Finita.

Quando $n = 1$, a Equação (1.26) reduz-se a

$$x + y = \binom{1}{0} x^0 y^1 + \binom{1}{1} x^1 y^0 = y + x.$$

Supondo a Equação (1.26) válida para $n - 1$. Então,

$$\begin{aligned}
(x+y)^n &= (x+y)(x+y)^{n-1} \\
&= (x+y) \sum_{i=0}^{n-1} \binom{n-1}{i} x^i y^{n-i-1} \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} x^{i+1} y^{n-i-1} + \sum_{i=0}^{n-1} \binom{n-1}{i} x^i y^{n-i}
\end{aligned}$$

Fazendo $k = i + 1$ no primeiro somatório e $k = i$ no segundo, temos

$$(x+y)^n = \sum_{k=1}^{n} \binom{n-1}{k-1} x^k y^{n-k} + \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-k}$$

$$
\begin{aligned}
&= x^n + \sum_{k=1}^{n-1}\left[\binom{n-1}{k-1} + \binom{n-1}{k}\right] x^k y^{n-k} + y^n \\
&= x^n + \sum_{k=1}^{n-1}\binom{n}{k} x^k y^{n-k} + y^n \\
&= \sum_{k=0}^{n}\binom{n}{k} x^k y^{n-k}.
\end{aligned}
$$

A penúltima igualdade segue da Equação (1.25).

Fazendo $x = p$ e $y = 1 - p$ na Equação (1.26), obtemos

$$
\sum_{i=0}^{n}\binom{n}{i} p^i (1-p)^{n-i} = (p+1-p)^n = 1. \tag{1.27}
$$

Logo, a soma da f.p. da distribuição binomial é igual a 1.

Podemos, então, calcular o valor esperado da distribuição binomial.

Da Definição 1.19 e da Equação (1.24), temos

$$
EX = \sum_{x=0}^{n} x\binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x\binom{n}{x} p^x (1-p)^{n-x}
$$

pois, em $x = 0$ o termo é 0.

Usando o fato que

$$
x\binom{n}{x} = x\frac{n!}{(n-x)!\,x!} = \frac{n(n-1)!}{(n-x)!\,(x-1)!} = n\binom{n-1}{x-1}.
$$

Temos,

$$
EX = \sum_{x=1}^{n} n\binom{n-1}{x-1} p^x (1-p)^{n-x}.
$$

Fazendo $x = y + 1$,

$$
EX = n\sum_{y=0}^{n-1}\binom{n-1}{y} p^{y+1}(1-p)^{n-(y+1)} = np\sum_{y=0}^{n-1}\binom{n-1}{y} p^{y}(1-p)^{n-y-1}.
$$

Como a última soma é a soma da f.p. da distribuição binomial $(n - 1, p)$ ela deve ser igual a 1.

Portanto,

$$
EX = np. \tag{1.28}
$$

Para calcular a variância da distribuição binomial, devemos encontrar o valor e EX^2. Temos,

$$EX^2 = \sum_{x=1}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$$

Usando o fato que

$$x^2 \binom{n}{x} = x^2 \frac{n!}{(n-x)!\,x!} = x \frac{n(n-1)!}{(n-x)!\,(x-1)!} = nx \binom{n-1}{x-1}.$$

Temos,

$$EX^2 = \sum_{x=1}^{n} nx \binom{n-1}{x-1} p^x (1-p)^{n-x}.$$

Fazendo $x = y + 1$,

$$\begin{aligned} EX^2 &= n \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^{y+1} (1-p)^{n-(y+1)} \\ &= np \sum_{y=0}^{n-1} y \binom{n-1}{y} p^y (1-p)^{n-y-1} + np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-y-1}. \end{aligned}$$

A primeira soma é o valor esperado de uma distribuição binomial $(n - 1, p)$, sendo igual a $(n - 1)p$ e a segunda soma é igual a 1. Então,

$$EX^2 = (n-1)np^2 + np. \tag{1.29}$$

Da Equação (1.18) e de (1.28) e (1.29), segue que

$$Var(X) = (n-1)np^2 + np - (np)^2 = np - np^2.$$

Logo,

$$Var(X) = np(1-p). \tag{1.30}$$

5.3 Distribuição de Poisson

A distribuição de Poisson é utilizada quando desejamos obter o número de ocorrências de um certo tipo de evento em um determinado intervalo de tempo.

Uma variável aleatória X com valores nos inteiros não negativos tem uma distribuição de Poisson com parâmetro $\lambda > 0$, a qual denotamos $X \sim Poisson(\lambda)$, se sua função de probabilidade é

$$P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \text{ com } x = 0, 1, \ldots. \tag{1.31}$$

Para calcularmos o valor esperado precisamos da Fórmula de Taylor, que aproxima uma função por um polinômio.

Definição 1.27. Seja $f : I \to \mathbb{R}$ uma função que admite derivadas até ordem r, isto é, $f^{(r)}(x) = \frac{d^r}{d_{x^r}} f(x)$ num ponto c do intervalo I. O polinômio de Taylor de ordem r de f no ponto c é dado por

$$T_r(x) = \sum_{i=0}^{r} \frac{f^{(i)}(c)}{i!}(x-c)^i.$$

Exemplo 1.1. Cálculo do polinômio de Taylor de $f(x) = e^x$ no ponto $c = 0$.

Temos,

$$f(x) = f'(x) = f''(x) = \cdots = f^{(n)}(x) = e^x$$

e

$$f(0) = f'(0) = f''(0) = \cdots = f^{(n)}(0) = e^0 = 1.$$

Logo,

$$e^x = \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!}(x-0)^i = \sum_{i=0}^{\infty} \frac{x^i}{i!}. \tag{1.32}$$

O valor esperado da distribuição de Poisson é dado por:

$$EX = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}.$$

Fazendo $y = x - 1$, obtemos

$$EX = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{(y)!}.$$

Da Equação (1.32), segue que

$$EX = \lambda e^{-\lambda} e^{\lambda}$$

e, portanto,

$$EX = \lambda. \tag{1.33}$$

Para determinar a variância, calculamos o valor de $EX2$.

$$EX^2 = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}.$$

Fazendo $y = x - 1$, obtemos

$$EX^2 = \lambda \sum_{y=0}^{\infty} (y+1) \frac{e^{-\lambda}\lambda^y}{y!} = \lambda \left[\sum_{y=0}^{\infty} y \frac{e^{-\lambda}\lambda^y}{y!} + \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} \right].$$

A primeira soma é o valor esperado e a segunda é igual a 1 pela Equação (1.32). Então,

$$EX^2 = \lambda(\lambda + 1). \tag{1.34}$$

Logo, da Equação (1.18) e de (1.33) e (1.34) temos

$$Var(X) = \lambda(\lambda + 1) - \lambda^2$$

e, portanto.

$$Var(x) = \lambda. \tag{1.35}$$

6 I DISTRIBUIÇÕES DE PROBABILIDADE CONTÍNUAS

6.1 Distribuição uniforme

Uma variável aleatória X é distribuída uniformemente ao longo do intervalo $[\alpha, \beta]$ se sua f.d.p. é dada por

$$f_X(x|\alpha,\beta) = \begin{cases} \frac{1}{\beta - \alpha}, & \text{se } \alpha \leqslant x \leqslant \beta; \\ 0, & \text{caso contrário.} \end{cases} \tag{1.36}$$

O valor esperado da distribuição uniforme é dado por:

$$\begin{aligned} EX &= \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx \\ &= \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} \\ &= \frac{\beta + \alpha}{2}. \end{aligned} \tag{1.37}$$

Para calcular a variância, primeiro calculamos EX^2.

$$\begin{aligned} EX^2 &= \int_{\alpha}^{\beta} \frac{x^2}{\beta-\alpha}dx \\ &= \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} \\ &= \frac{\beta^2+\alpha\beta+\alpha^2}{3}. \end{aligned} \tag{1.38}$$

Da Equação (1.18) e de (1.37) e (1.38), obtemos:

$$Var(X) = \frac{\beta^2+\alpha\beta+\alpha^2}{3} - \frac{(\beta+\alpha)^2}{4}.$$

Portanto,

$$Var(X) = \frac{(\beta-\alpha)^2}{12}. \tag{1.39}$$

6.2 Distribuição normal

Uma variável aleatória X tem distribuição normal, com parâmetros μ, $-\infty < \mu < \infty$ e $\sigma > 0$, denotada por $X \sim N(\mu, \sigma^2)$, se sua f.d.p. é dada por

$$f_X(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < \infty. \tag{1.40}$$

Para $\mu = 0$ e $\sigma = 1$, a variável aleatória $Z = \frac{(X-\mu)}{\sigma}$ tem uma distribuição $Z \sim N(0, 1)$, conhecida como distribuição normal padrão.

Fazendo $z = \frac{(x-\mu)}{\sigma}$, a f.d.p. reduz-se a

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}, \ -\infty < z < \infty. \tag{1.41}$$

Para provar que 1.40 é de fato uma f.d.p., precisamos mostrar que

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 1 \tag{1.42}$$

ao longo de toda reta real.

A demonstração de (1.42) foi extraída de [1].

Fazendo $z = \frac{(x-\mu)}{\sigma}$, obtemos

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz.$$

Como a integral sobre $(-\infty, 0)$ é igual a integral sobre $(0, \infty)$, pois na distribuição

normal padrão obtemos um gráfico simétrico, vide Figura 1, reduzimos o problema mostrando que

$$\int_0^\infty e^{\frac{-z^2}{2}} dz = \sqrt{\frac{\pi}{2}}.$$

Como ambos os lados da igualdade são positivos, a igualdade se manterá admitindo que os quadrados são iguais, então

$$\begin{aligned} \left(\int_0^\infty e^{\frac{-z^2}{2}} dz\right)^2 &= \left(\int_0^\infty e^{\frac{-u^2}{2}} du\right)\left(\int_0^\infty e^{\frac{-v^2}{2}} dv\right) \\ &= \int_0^\infty \int_0^\infty e^{\frac{-(u^2+v^2)}{2}} dudv. \end{aligned}$$

Fazendo uma mudança de variáveis para coordenadas polares, obtemos $u = r\cos\theta$ e $v = r\,\text{sen}\theta$.

Dessa forma, $u^2 + v^2 = r^2$, $dudv = rd\theta dr$ e os limites de integração passam a ser $0 < r < \infty$ e $0 < \theta < \frac{\pi}{2}$. Então,

$$\begin{aligned} \left(\int_0^\infty e^{\frac{-z^2}{2}} dz\right)^2 &= \int_0^\infty \int_0^{\frac{\pi}{2}} e^{\frac{-r^2}{2}} rd\theta dr \\ &= \frac{\pi}{2} \int_0^\infty re^{\frac{-r^2}{2}} dr \\ &= \frac{\pi}{2}\left[-e^{\frac{-r^2}{2}} \Big|_0^\infty\right] \\ &= \frac{\pi}{2}, \end{aligned}$$

o que demonstra (1.42).

Para demonstrar o valor esperado e a variância da distribuição normal, precisamos de dois teoremas.

Teorema 1.28. *Sejam X uma variável aleatória, e a, b e c constantes. Para quaisquer funções* $g_1(X)$ *e* $g_2(X)$*, desde que seus valores esperados existam, temos*

$$E(ag_1(X) + bg_2(X) + c) = aEg_1(X) + bEg_2(X) + c.$$

Demonstração. No caso das variáveis aleatórias discretas, temos

$$\begin{aligned} E(ag_1(X) + bg_2(X) + c) &= \sum (ag_1(X) + bg_2(X) + c)\, f_X(x) \\ &= \sum ag_1(X)\, f_X(x) + \sum bg_2(X)\, f_X(x) + \sum c\, f_X(x) \\ &= a\sum g_1(X)\, f_X(x) + b\sum g_2(X)\, f_X(x) + c\sum f_X(x) \\ &= aEg_1(X) + bEg_2(X) + c. \end{aligned}$$

No caso em que as variáveis aleatórias são contínuas,

$$
\begin{aligned}
E(ag_1(X)+bg_2(X)+c) &= \int_{-\infty}^{\infty}(ag_1(X)+bg_2(X)+c)f_X(x)dx\\
&= \int_{-\infty}^{\infty}ag_1(X)f_X(x)dx+\int_{-\infty}^{\infty}bg_2(X)f_X(x)dx+\int_{-\infty}^{\infty}cf_X(x)dx\\
&= a\int_{-\infty}^{\infty}g_1(X)f_X(x)dx+b\int_{-\infty}^{\infty}g_2(X)f_X(x)dx+c\int_{-\infty}^{\infty}f_X(x)dx\\
&= aEg_1(X)+bEg_2(X)+c.
\end{aligned}
$$

Teorema 1.29. *Seja X uma variável aleatória com variância finita, então, para quaisquer a e b constantes, temos*

$$Var(aX+b)=a^2Var(X).$$

Demonstração. Da Equação (1.15) e aplicando o Teorema 1.28 duas vezes, obtemos

$$
\begin{aligned}
Var(aX+b) &= E[(aX+b)-E(aX+b)]^2\\
&= E[(aX+b)-aEX+b]^2\\
&= E(aX-aEX)^2\\
&= E[a^2(X-EX)^2]\\
&= a^2E(X-EX)^2\\
&= a^2Var(X).
\end{aligned}
$$

Inicialmente, determinamos o valor esperado e a variância de uma variável aleatória padrão $Z=\frac{(X-\mu)}{\sigma}$.

$$EZ=\int_{-\infty}^{\infty}zf_Z(z)dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{\frac{-z2}{2}}dz=-\frac{1}{\sqrt{2\pi}}e^{\frac{-z2}{2}}|_{-\infty}^{\infty}=0.$$

Dessa forma temos,

$$Var(Z)=EZ^2=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{\frac{-z2}{2}}dz.$$

Fazendo $u=z$ e $dv=ze^{\frac{-z^2}{2}}$ e integrando por partes, obtemos

$$
\begin{aligned}
Var(Z) &= \frac{1}{\sqrt{2\pi}}\left[-e^{\frac{-z2}{2}}|_{-\infty}^{\infty}+\int_{-\infty}^{\infty}e^{\frac{-z2}{2}}dz\right]\\
&= -\frac{1}{\sqrt{2\pi}}e^{\frac{-z2}{2}}|_{-\infty}^{\infty}+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{\frac{-z2}{2}}dz,
\end{aligned}
$$

onde a primeira parcela é igual a $EZ=0$ e a segunda, pela Equação (1.42), é igual a 1.

Portanto,

$$Var(Z) = 1.$$

Para $X \sim N(\mu, \sigma^2)$, temos, pelo Teorema 1.28,

$$EX = E(\mu + \sigma Z) = \mu + \sigma EZ = \mu. \qquad (1.43)$$

Da mesma forma, pelo Teorema 1.29,

$$Var(X) = Var(\mu + \sigma Z) = \sigma^2 Var(Z) = \sigma^2. \qquad (1.44)$$

O desvio padrão é a raiz quadrada da variância.

$$DP(X) = \sigma. \qquad (1.45)$$

A curva da distribuição normal é simétrica, com valor máximo em $x = \mu$ e pontos de inflexão em $\mu \pm \sigma$.

De (1.42), a área sob a curva normal é igual a 1. A probabilidade dentro de 1 desvio padrão é:

$$\begin{aligned} P(\mu - \sigma \leq X \leq \mu + \sigma) &= P(-1 \leq Z \leq 1) \\ &= \Phi(1) - \Phi(-1) \\ &= \Phi(1) - [1 - \Phi(1)] \\ &= 0,8413 - [1 - 0,8413] \\ &= 0,6826. \end{aligned}$$

O valor de $\Phi(1)$ foi obtido da Tabela 2.

Observação 1.6.1. $\Phi(-x) = 1 - \Phi(x)$.

Procedendo de forma análoga, obtemos:

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq Z \leq 2) = 0,9544$$

e

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq 3) = 0,9974.$$
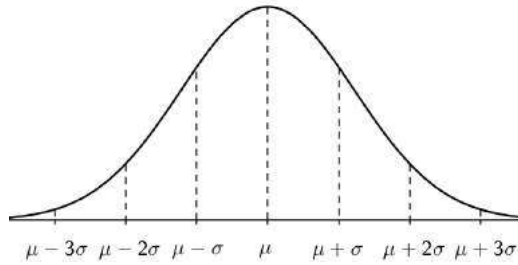
Figura 1: Curva normal com desvios padrão

Fonte: Arquivo próprio do autor, baseado em [10].

6.3 Distribuição exponencial

Uma variável aleatória X tem distribuição exponencial, com parâmetro $\lambda > 0$, denotada por $X \sim Exp(\lambda)$, se sua f.d.p. é dada por

$$f_X(x|\lambda) = \begin{cases} \lambda e^{-\lambda x}, & \text{se } x \geqslant 0; \\ 0, & \text{se } x < 0. \end{cases} \tag{1.46}$$

O valor esperado de uma distribuição exponencial é

$$EX = \int_0^\infty x\lambda e^{-\lambda x} dx.$$

Vamos calcular EX^n para $n > 0$.

O cálculo de EX^n foi extraído de [15].

$$EX^n = \int_0^\infty x^n \lambda e^{-\lambda x} dx.$$

Fazendo $u = x^n$ e $dv = \lambda e^{-\lambda x}$ e integrando por partes, obtemos

$$\begin{aligned} EX^n &= -x^n e^{-\lambda x}|_0^\infty + n\int_0^\infty x^{n-1} e^{-\lambda x} dx \\ &= \frac{n}{\lambda}\int_0^\infty \lambda x^{n-1} e^{-\lambda x} dx \\ &= \frac{n}{\lambda} EX^{n-1}. \end{aligned}$$

Para $n = 1$, temos

$$EX = \frac{1}{\lambda}. \tag{1.47}$$

Para $n = 2$, temos

$$EX^2 = \frac{2}{\lambda}EX = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2}. \tag{1.48}$$

Pela Equação (1.18) e de 1.47 e 1.48, a variância da distribuição exponencial é dada por

$$Var(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}. \tag{1.49}$$

6.4 Distribuição gama

Uma variável aleatória X tem distribuição gama, com parâmetros $\alpha > 0$ e $\lambda > 0$, se sua f.d.p. é dada por

$$f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & \text{se } x \geqslant 0; \\ 0, & \text{se } x < 0. \end{cases} \tag{1.50}$$

onde definimos a função gama, denominada por $\Gamma(\alpha)$, como

$$\Gamma(\alpha) = \int_0^\infty e^{-y}y^{\alpha-1}dy. \tag{1.51}$$

A média de uma distribuição gama é dada por

$$EX = \frac{1}{\Gamma(\alpha)}\int_0^\infty \lambda x e^{-\lambda x}(\lambda x)^{\alpha-1}dx = \frac{1}{\Gamma(\alpha)}\int_0^\infty e^{-\lambda x}(\lambda x)^{\alpha}dx.$$

Fazendo $u = \lambda x$, obtemos $du = \lambda dx$. Então,

$$EX = \frac{1}{\lambda\Gamma(\alpha)}\int_0^\infty e^{-u}u^{\alpha}du = \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)}.$$

Integrando $\Gamma(\alpha + 1) = \int_0^\infty e^{-y}y^{\alpha}dy$ por partes, obtemos

$$\begin{aligned} \Gamma(\alpha+1) &= -e^{-y}y^{\alpha}\big|_0^\infty + \int_0^\infty e^{-y}\alpha y^{\alpha-1}dy \\ &= \alpha\int_0^\infty e^{-y}y^{\alpha-1}dy \\ &= \alpha\Gamma(\alpha). \end{aligned}$$

Portanto,

$$EX = \frac{\alpha}{\lambda}. \tag{1.52}$$

Para encontrar $Var(x)$, calculemos EX^2.

$$EX^2 = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda x^2 e^{-\lambda x} (\lambda x)^{\alpha-1} dx = \frac{1}{\lambda\Gamma(\alpha)} \int_0^\infty e^{-\lambda x} (\lambda x)^{\alpha+1} dx.$$

Fazendo $u = \lambda x$, obtemos $du = \lambda dx$. Então,

$$EX^2 = \frac{1}{\lambda^2\Gamma(\alpha)} \int_0^\infty e^{-u} u^{\alpha+1} du = \frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)}.$$

Integrando $\Gamma(\alpha + 2) = \int_0^\infty e^{-y} y^{\alpha+1} dy$ por partes, obtemos

$$\begin{aligned} \Gamma(\alpha+2) &= -e^{-y} y^{\alpha+1} \Big|_0^\infty + \int_0^\infty e^{-y} (\alpha+1) y^\alpha dy \\ &= (\alpha+1) \int_0^\infty e^{-y} y^\alpha dy \\ &= (\alpha+1)\Gamma(\alpha+1) \\ &= \alpha(\alpha+1)\Gamma(\alpha). \end{aligned}$$

Portanto,

$$EX^2 = \frac{\alpha(\alpha+1)}{\lambda^2}. \tag{1.53}$$

Pela Equação (1.18), segue que

$$Var(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}. \tag{1.54}$$

Distribuição qui-quadrada

Definição 1.30. Para variáveis aleatórias normal padrão independentes $Z_1, \ldots, Z_n$, chamamos $\chi_n^2 = Z_1^2 + \cdots + Z_n^2$ de variável aleatória qui-quadrada com n graus de liberdade.

Seja n um inteiro positivo, a distribuição gama com $\lambda = 1/2$ e $\alpha = n/2$, é chamada de distribuição qui-quadrada com n graus de liberdade.

$$f_{\chi_n^2}(x) = \begin{cases} \dfrac{x^{[(n/2)-1]} e^{-x/2}}{2^{n/2}\Gamma(n/2)}, & \text{se } x \geqslant 0; \\ 0, & \text{se } x < 0. \end{cases} \tag{1.55}$$

De (1.52) e (1.54), obtemos:

$$E(\chi_n^2) = n \tag{1.56}$$

e

$$Var(\chi_n^2) = 2n. \tag{1.57}$$

7 | VARIÁVEIS ALEATÓRIAS MÚLTIPLAS

Em um experimento, frequentemente nos deparamos com mais de uma variável aleatória. Inicialmente, nos ocuparemos com o caso de duas variáveis aleatórias.

Definição 1.31. Para quaisquer duas variáveis aleatórias discretas X e Y, definimos a função $f_{X,Y}(x, y) : \mathbb{R}^2 \to \mathbb{R}$, chamada função de probabilidade conjunta, ou simples- mente, f.p. conjunta de X e Y como

$$f_{X,Y}(x, y) = P(X = x,\ Y = y).$$

As funções de probabilidade marginal de X e Y, obtidas de $f_{X,Y}(x, y)$, são dadas por

$$f_X(x) = P(X = x) = \sum_y f_{X,Y}(x, y).$$

$$f_Y(y) = P(Y = y) = \sum_x f_{X,Y}(x, y).$$

Definição 1.32. Para quaisquer duas variáveis aleatórias contínuas X e Y, definimos a função $f_{X,Y}(x, y) : \mathbb{R}^2 \to \mathbb{R}$, chamada função densidade de probabilidade conjunta, ou simplesmente, f.d.p. conjunta de X e Y, para cada $A \in \mathbb{R}$ como

$$P((X,\ Y) \in A) = \int_A\int f_{X,Y}(x, y)dxdy.$$

As funções densidade de probabilidade marginal de X e Y, obtidas de $f_{X,Y}(x, y)$, são dadas por

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy, \text{ com } -\infty < x < \infty.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx, \text{ com } -\infty < y < \infty.$$

Definição 1.33. O valor esperado de uma variável aleatória, definida pela função $g(X,\ Y)$, representado por $Eg(X,\ Y)$ é

$$Eg(X,Y) = \begin{cases} \sum_{(x,y)} g(X,Y)f_{X,Y}(x,y), & \text{no caso discreto;} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X,Y)f_{X,Y}(x,y)dxdy, & \text{no caso contínuo.} \end{cases}$$

7.1 Distribuições condicionais e independência

Para duas variáveis aleatórias X e Y, com f.p. ou f.d.p. conjunta $f(x, y)$ e f.p. ou f.d.p. marginal $f_X(x)$ e $f_Y(y)$, temos as seguintes definições:

Definição 1.34. Para qualquer x tal que $P(X = x) = f_X(x) > 0$, definimos a f.p., ou

f.d.p. condicional de Y dado que $X = x$ como a função de y por

$$f(y|x) = \frac{f(x,y)}{f_X(x)}.$$

Para qualquer y tal que $P(Y = y) = f_Y(y) > 0$, definimos a f.p., ou f.d.p. condicional de X dado que $Y = y$ como a função de y por

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}.$$

Definição 1.35. O valor esperado condicional de uma função de uma variável aleatória $g(X)$, dado que $Y = y$, denotado por $E(g(X)|Y)$ é

$$E(g(X)|Y) = \begin{cases} \sum_x g(X) f(x|y), & \text{no caso discreto;} \\ \int_{-\infty}^{\infty} g(X) f(x|y) dx, & \text{no caso contínuo.} \end{cases}$$

Para $E(g(Y)|X)$ temos uma definição análoga.

O Teorema 1.28 é válido para $Eg(X, Y)$, $E(g(X)|Y)$ e $E(g(Y)|X)$.

Definição 1.36. As variáveis aleatórias X e Y, para cada $x, y \in \mathbb{R}$, são variáveis aleatórias independentes se

$$f(x, y) = f_X(x)\, f_Y(y).$$

Se X e Y são independentes, temos, então,

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f_X(x) f_Y(y)}{f_Y(y)} = f_X(x)$$

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y).$$

Teorema 1.37. *Sejam X e Y variáveis aleatórias, tais que EX e EY sejam ambos finitos, então*

$$E(X + Y) = EX + EY.$$

Demonstração. Tome na Definição 1.33 $g(X, Y) = X + Y$.

Para o caso discreto, temos:

$$
\begin{aligned}
E(X+Y) &= \sum_x \sum_y (x+y) f(x,y) \\
&= \sum_x \sum_y x f(x,y) + \sum_x \sum_y y f(x,y) \\
&= \sum_x x \sum_y f(x,y) + \sum_y y \sum_x f(x,y) \\
&= \sum_x x f_X(x) + \sum_y y f_Y(y) \\
&= EX + EY.
\end{aligned}
$$

Para o caso contínuo, temos:

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f(x,y) dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y) dydx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y) dxdy \\
&= \int_{-\infty}^{\infty} x f_X(x) dx + \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= EX + EY.
\end{aligned}
$$

Teorema 1.38. *Sejam X e Y variáveis aleatórias independentes, então*

$$E(XY) = EX \cdot EY.$$

Demonstração. Tome na Definição 1.33 $g(X, Y) = XY$.

Para o caso discreto, temos:

$$
\begin{aligned}
E(XY) &= \sum_x \sum_y xy f(x,y) \\
&= \sum_x \sum_y xy f_X(x) f_Y(y) \\
&= \sum_x x f_X(x) \sum_y y f_Y(y) \\
&= EX \cdot EY.
\end{aligned}
$$

Para o caso contínuo, temos:

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x,y) dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dxdy
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} y f_Y(y) \int_{-\infty}^{\infty} x f_X(x) dx dy \\
&= \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy \\
&= EX \cdot EY.
\end{aligned}
$$

Podemos generalizar o Teorema 1.38.

Considere as variáveis aleatórias $X_1, \ldots, X_n$, mutuamente independentes. Então,

$$E(X_1, \ldots, X_n) = EX_1 \cdot \cdots \cdot EX_n. \quad (1.58)$$

7.2 Covariância e correlação

A covariância é uma medida da relação linear entre as variáveis aleatórias X e Y.

Definição 1.39. A covariância de duas variáveis aleatórias é um número definido por

$$Cov(X, Y) = E[(X - EX)(Y - EY)].$$

Podemos simplificar o cálculo da covariância.

$$
\begin{aligned}
Cov(X, Y) &= E[(X - EX)(Y - EY)] \\
&= E[XY - XEY - YEX + EX \cdot EY] \\
&= E(XY) - EX \cdot EY - EY \cdot EX + EX \cdot EY \\
&= E(XY) - EX \cdot EY. \quad (1.59)
\end{aligned}
$$

A correlação é uma medida que não depende das unidades de medida das variáveis aleatórias X e Y.

Definição 1.40. A correlação de duas variáveis aleatórias é um número definido por

$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}.$$

Teorema 1.41. *Se as variáveis aleatórias X e Y são independentes, então $Cov(X, Y) = 0$ e $\rho_{XY} = 0$.*

Demonstração. Pela Teorema 1.38 temos

$$E(XY) = EX \cdot EY$$

e, portanto, da Equação (1.59), obtemos

$$Cov(X, Y) = EX \cdot EY - EX \cdot EY = 0$$

e

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0.$$

Teorema 1.42. *Sejam X e Y duas variáveis aleatórias, e a e b constantes, então*

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y).$$

Demonstração. A demonstração foi extraída de [1].

Do Teorema 1.28 temos

$$E(aX + bY) = aEX + bEY.$$

Logo, pela Definição 1.15,

$$\begin{aligned} Var(aX + bY) &= E[(aX + bY) - (aEX + bEY)]^2 \\ &= E[a(X - EX) + b(Y - EY)]^2 \\ &= E[a^2(X - EX)^2 + b^2(Y - EY)^2 + 2ab(X - EX)(Y - EY)] \\ &= a^2E(X - EX)^2 + b^2E(Y - EY)^2 + 2abE[(X - EX)(Y - EY)] \\ &= a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y). \end{aligned}$$

Se as variáveis aleatórias X e Y são independentes, então $Cov(X, Y) = 0$ e, portanto,

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y).$$

8 I FUNÇÕES GERATRIZES DE MOMENTOS

A função geratriz de momentos (f.g.m.) é uma função associada à distribuição de probabilidade, auxiliando a caracterizá-la.

Definição 1.43. A função geratriz de momentos de uma variável aleatória X, denotada por $M_X(t)$, para todos os valores reais de t, como

$$M_X(t) = Ee^{tX} = \begin{cases} \sum_x e^{tX} P(X = x), & \text{se } X \text{ for discreta;} \\ \int_{-\infty}^{\infty} e^{tX} f_X(x) dx, & \text{se } X \text{ for contínua.} \end{cases}$$

Os momentos da variável aleatória X podem ser obtidos das sucessivas derivadas de $M_X(t)$ avaliada em $t = 0$, como no teorema a seguir.

Teorema 1.44. *Se X á uma variável aleatória e possui um f.g.m.* $M_X(t)$, *então*

$$EX^n = M_X^{(n)}(0). \tag{1.60}$$

Definimos

$$M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t)|_{t=0}.$$

Demonstração. No caso discreto, consideramos que

$$\frac{d}{dt}\sum_x e^{tx}P(X=x) = \sum_x \left(\frac{d}{dt}\right) e^{tx}P(X=x).$$

Logo,

$$\frac{d}{dt}M_X(t) = \frac{d}{dt}\sum_x e^{tx}P(X=x) = \sum_x \left(\frac{d}{dt}\right) e^{tx}P(X=x) = \sum_x \left(xe^{tx}\right) P(X=x) = EXe^{tx}.$$

No caso contínuo, assumindo como legítima a hipótese de diferenciar sob o sinal da integral, temos

$$\frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f_X(x)dx = \int_{-\infty}^{\infty} \left(\frac{d}{dt}\right) e^{tx} f_X(x)dx.$$

Então,

$$\frac{d}{dt}M_X(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f_X(x)dx = \int_{-\infty}^{\infty} \left(\frac{d}{dt}\right) e^{tx} f_X(x)dx = \int_{-\infty}^{\infty} \left(xe^{tx}\right) f_X(x)dx = EXe^{tx}.$$

Portanto,

$$\frac{d}{dt}M_X(t)|_{t=o} = EXe^{tx}|_{t=o} = EX.$$

Analogamente, temos

$$\frac{d^n}{dt^n}M_X(t)|_{t=o} = EX^n e^{tx}|_{t=o} = EX^n.$$

Teorema 1.45. *Para a e b constantes, a f.g.m. da variável aleatória aX + b é dada por*

$$M_{aX+b}(t) = e^{bt}\, M_X(at).$$

Demonstração.

$$M_{aX+b}(t) = E(e^{(aX+b)t} = E(e^{aXt}e^{bt}).$$

Como e^{bt} é uma constante, segue do Teorema 1.28, que

$$M_{aX+b}(t) = e^{bt}E(e^{aXt}).$$

Da definição de f.g.m., temos

$$M_{aX+b}(t) = e^{bt}\, M_X(at).$$

A seguir são dados exemplos da f.g.m. de algumas distribuições.

Exemplo 1.2. F.g.m. da distribuição binomial.

Seja $X \sim Binomial(n, p)$, sua f.p. é dada por

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

Então,

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1 - p)^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1 - p)^{n-x}.$$

Do Teorema 1.26, obtemos

$$M_X(t) = [pe^t + (1 - p)]^n.$$

Exemplo 1.3. F.g.m. da distribuição de Poisson.

Seja $X \sim Poisson\ (\lambda)$, sua f.p. é dada por

$$P(X = x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}.$$

Então,

$$M_X(t) = \sum_{x=0}^{n} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{n} \frac{(\lambda e^t)^x}{x!}.$$

De 1.32 temos,

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}.$$

Portanto,

$$M_X(t) = e^{-\lambda} e^{\lambda e^t} = exp\ \ \lambda(e^t - 1)\}\ .$$

Exemplo 1.4. F.g.m. da distribuição normal.

Seja $Z \sim N(0, 1)$, sua f.d.p. é dada por

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}.$$

Então,

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{\frac{-z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left\{-\frac{z^2}{2} + tz\right\} dz.$$

Adicionando e subtraindo $\frac{t^2}{2}$, obtemos,

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left\{-\frac{z^2}{2} + tz - \frac{t^2}{2} + \frac{t^2}{2}\right\} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left\{-\frac{(z-t)^2}{2} + \frac{t^2}{2}\right\} dz.$$

Logo,

$$M_Z(t) = e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2}} dz.$$

Como, de (1.42)

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = 1.$$

Segue que,

$$M_Z(t) = e^{\frac{t^2}{2}}. \quad (1.61)$$

Para obtermos $X \sim N(\mu, \sigma^2)$ basta fazer $X = \mu + \sigma Z$, então,

$$\begin{aligned} M_X(t) &= Ee^{tX} \\ &= Ee^{t(\mu+\sigma Z)} \\ &= E\left(e^{t\mu} e^{t\sigma Z}\right) \end{aligned}$$

Do Teorema 1.45, segue que

$$\begin{aligned} M_X(t) &= e^{t\mu} E\left(e^{t\sigma Z}\right) \\ &= e^{t\mu} M_Z(t\sigma) \\ &= e^{t\mu} e^{\frac{(t\sigma)^2}{2}} \\ &= exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}. \end{aligned}$$

Exemplo 1.5. F.g.m. da distribuição exponencial.

Seja $X \sim Exp(\lambda)$, sua f.d.p. é dada por

$$f_X(x|\lambda) = \lambda e^{-\lambda x}, \text{ para todo } x \geq 0.$$

Então,

$$M_X(t) = \lambda \int_0^\infty e^{tx} e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t} e^{-(\lambda-t)x} \Big|_0^\infty .$$

Portanto,

$$M_X(t) = \frac{\lambda}{\lambda - t}, \text{ para } t < \lambda \ .$$

Exemplo 1.6. F.g.m. da distribuição gama.

Seja X uma variável aleatória com distribuição gama, sua f.d.p. é dada por

$$f_X(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha - 1}}{\Gamma(\alpha)}, \text{ para todo } x \geqslant 0.$$

Então,

$$M_X(t) = \frac{\lambda}{\Gamma(\alpha)} \int_0^\infty e^{tx} e^{-\lambda x} (\lambda x)^{\alpha-1} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-(\lambda-t)x} x^{\alpha-1} dx.$$

Fazendo $u = (\lambda - t)x$, obtemos $du = (\lambda - t)dx$, então,

$$M_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-u} \left(\frac{u}{\lambda - t} \right)^{\alpha - 1} \frac{du}{\lambda - t} = \left(\frac{\lambda}{\lambda - t} \right)^\alpha \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-u} u^{\alpha-1} du.$$

De (1.51), temos,

$$\Gamma(\alpha) = \int_0^\infty e^{-u} u^{\alpha - 1} du.$$

Portanto,

$$M_X(t) = \left(\frac{\lambda}{\lambda - t} \right)^\alpha, \text{ para } t < \lambda \ .$$

Teorema 1.46. *Sejam X e Y variáveis aleatórias independentes, que possuam f.g.m. $M_X(t)$ e $M_Y(t)$, respectivamente, tais que $Z = X + Y$. Então,*

$$M_Z(t) = M_X(t)M_Y(t).$$

Demonstração.

$$M_Z(t) = E(e^{t(X+Y)} = E(e^{tX}e^{tY}).$$

Do Teorema 1.38, temos,

$$M_Z(t) = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t).$$

Podemos generalizar o Teorema 1.46.

Considere as variáveis aleatórias $X_1, \ldots, X_n$, mutuamente independentes, tais que $Z = X_1 + \cdots + X_n$. Então,

$$M_Z(t) = M_X1\ (t) \cdot \cdots \cdot M_Xn\ (t). \quad (1.62)$$

Em particular, se todos os $X_1, \ldots, X_n$ têm a mesma distribuição com f.g.m. $M_X(t)$. Então,

$$M_Z(t) = [M_X(t)]^n. \quad (1.63)$$

INFERÊNCIA ESTATÍSTICA

O ramo da estatística que faz uso de uma amostra para fazer afirmações sobre carac- terísticas de uma população, tendo como ferramenta a probabilidade, é o da inferência estatística.

1 I CONCEITOS INICIAIS

1.1 População e amostra

A seguir serão apresentadas as definições de população e amostra, importantes no desenvolvimento deste trabalho.

Definição 2.1. Uma população é um conjunto de todos os elementos ou resultados de uma característica de interesse. Um subconjunto de uma população é dita uma amostra.

1.2 Amostras aleatórias simples

A amostragem aleatória simples (a.a.s.) é um modo para selecionar uma amostra de uma população, de tal forma que todos os elementos da população, em um processo aleatório de seleção, tenham a mesma probabilidade de serem selecionados.

No decorrer deste trabalho, serão consideradas as amostragens aleatórias simples com reposição, ou seja, uma amostra retirada tem a possibilidade de ser sorteada novamente, de modo que se tenha uma independência entre os elementos selecionados.

Definição 2.2. As variáveis aleatórias $X_1, \dots, X_n$ são chamadas amostra aleatória sim- ples de tamanho n de uma variável aleatória X se $X_1, \dots, X_n$ forem variáveis aleatórias independentes e identicamente distribuídas (i.i.d.).

1.3 Diferenciando estatística de parâmetro

Obtida uma amostra $X_1, \dots, X_n$, geralmente calculamos uma função $T(X_1, \dots, X_n)$ cujo domínio inclui todo o espaço amostral de $(X_1, \dots, X_n)$.

Definição 2.3. Dada uma amostra aleatória de tamanho n, $X_1, \dots, X_n$, uma estatística é uma variável aleatória, ou um vetor aleatório $Y = T(X_1, \dots, X_n)$, sendo $T(X_1, \dots, X_n)$ uma função com valor real ou por vetor, cujo domínio inclui todo o espaço amostral de $(X_1, \dots, X_n)$.

A variável aleatória, ou vetor aleatório $Y = T(X_1, \dots, X_n)$ descreve uma característica da amostra.

As estatísticas mais utilizadas são: a média amostral, a variância amostral e o desvio padrão amostral, definidas a seguir.

Definição 2.4. A média amostral é a média aritmética dos valores em uma variável aleatória.

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i, \text{ com } i = 1, \ldots, n.$$

Definição 2.5. A variância amostral é definida por

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \text{ com } i = 1, \ldots, n.$$

O desvio padrão amostral é definido por $S = \sqrt{S^2}$.

Definição 2.6. Um parâmetro descreve uma característica populacional.

Os parâmetros média e variância, de uma variável aleatória X, serão denotados, respectivamente, por EX e $Var(X)$.

2 I DISTRIBUIÇÕES AMOSTRAIS

Dada uma população X, estamos interessados em se afirmar algo sobre um parâmetro θ. Retiramos, então, amostras $X_1, \ldots, X_n$ de tamanho n por meio de uma determi- nada técnica de amostragem.

Para cada amostra, obtemos um valor t para a estatística $T = f(X_1, \ldots, X_n)$, que é uma função da amostra. Todos os valores de t formam uma nova população, e baseados nesta nova população que fazemos afirmações sobre o parâmetro θ.

A distribuição de T quando $(X_1, \ldots, X_n)$ assume todos os valores possíveis é chamada distribuição amostral da estatística T.

2.1 Distribuição amostral da média

Para calcular o valor esperado e a variância da média, precisamos do Lema a seguir.

Lema 2.7. *Sejam as variáveis aleatórias $X_1, \ldots, X_n$ de uma amostra aleatória simples de tamanho n de uma variável aleatória X e $g(X)$ uma função tal que existam $Eg(X_i)$ e $Varg(X_i)$, com $i = 1, \ldots, n$. Então,*

$$E\left(\sum_{i=1}^{n} g(X_i)\right) = n\left(Eg(X_1)\right) \tag{2.1}$$

e

$$Var\left(\sum_{i=1}^{n} g(X_i)\right) = n\left(Varg(X_1)\right). \tag{2.2}$$

Demonstração. Inicialmente provaremos (2.1).

Como $X_1, \ldots, X_n$ são variáveis aleatórias independentes e identicamente distribuídas, então $X_1 = \cdots = X_n$, logo $X_i = X_1$, de onde segue que

$$\sum_{i=1}^{n} g(X_i) = ng(X_1).$$

Portanto,

$$E\left(\sum_{i=1}^{n} g(X_i)\right) = E\left(ng(X_1)\right) = n\left(Eg(X_1)\right)$$

Para provar (2.2), temos, pela Equação (1.15),

$$Var\left(\sum_{i=1}^{n} g(X_i)\right) = E\left[\sum_{i=1}^{n} g(X_i) - E\left(\sum_{i=1}^{n} g(X_i)\right)\right]^2 = E\left[\sum_{i=1}^{n} (g(X_i) - Eg(X_i))\right]^2.$$

Temos, na última igualdade $n2$ termos. Sendo n termos $[g(X_i) - Eg(X_i)]^2$ e $n(n-1)$ termos $[g(X_i) - Eg(X_i)][g(X_j) - Eg(X_j)]$, com $i \neq j$.

Pela Definição 1.39 e pelo o fato de $X_1, \ldots, X_n$ serem variáveis aleatórias independentes, temos

$$E[(g(X_i) - Eg(X_i))(g(X_j) - Eg(X_j))] = Cov(X_i, X_j) = 0.$$

Pela Equação (1.15) e pelo o fato de $X_1, \ldots, X_n$ serem variáveis aleatórias identicamente distribuídas, temos

$$E[g(X_i) - Eg(X_i)]^2 = Var(X_i) = Var(X_1).$$

Obtendo, assim, a Equação (2.2).

Teorema 2.8. *Sejam $X_1, \ldots, X_n$ uma amostra aleatória de uma população com $EX = \mu$ e $Var(X) = \sigma^2 < \infty$. Temos, então,*

$$E\bar{X} = \mu \tag{2.3}$$

e

$$Var(\bar{X}) = \frac{\sigma^2}{n} \tag{2.4}$$

Demonstração. Pela Definição 2.4, temos

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Para provar (2.3) temos, pelo Teorema 1.28 e pela Equação (2.1),

$$E\bar{X} = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}nEX_1 = \mu.$$

Para provar (2.4) temos, pelo Teorema 1.29 e pela Equação (2.2),

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}nVar(X_1) = \frac{\sigma^2}{n}.$$

O desvio padrão da distribuição amostral da média, ou erro padrão da média, denotado por $\sigma_{\bar{X}}$, é

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}. \quad (2.5)$$

2.2 Distribuição normal da média

Para encontrar a distribuição normal da média, o Teorema a seguir e as funções geratrizes de momentos mostram-se úteis.

Teorema 2.9. *Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma população com média $\bar{X} = \frac{1}{n} X_1 + \cdots + X_n$ e f.g.m. $M_X(t)$. Então,*

$$M_{\bar{X}}(t) = [M_X(t/n)]^n.$$

Demonstração. Considere $Y = X_1 + \cdots + X_n$, então pela Definição 1.43 temos,

$$M_{\bar{X}}(t) = Ee^{t\bar{X}} = Ee^{t\left(\frac{1}{n}\right)(X_1+\cdots+X_n)} = Ee^{\frac{t}{n}Y} = M_Y(t/n).$$

Sendo $X_1, \ldots, X_n$ identicamente distribuídas, então,

$$Y = nX_1 = nX.$$

E pela Equação (1.63) temos,

$$M_{\bar{X}}(t) = [M_X(t/n)]^n.$$

Teorema 2.10. *Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma população normal $N(\mu, \sigma^2)$. Então,*

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right). \quad (2.6)$$

Demonstração. Do Exemplo 1.4 e do Teorema 2.9 temos,

$$
\begin{aligned}
M_{\bar{X}}(t) &= \left\{ exp\left\{ \mu\frac{t}{n} + \frac{\sigma^2\left(\frac{t}{n}\right)^2}{2} \right\} \right\}^n \\
&= exp\left\{ n\left\{ \mu\frac{t}{n} + \frac{\sigma^2\left(\frac{t}{n}\right)^2}{2} \right\} \right\} \\
&= exp\left\{ \mu t + \frac{\sigma^2 t^2}{2n} \right\}.
\end{aligned}
$$

Calculando a primeira e a segunda derivadas, obtemos,

$$M'_{\bar{X}}(t) = \left(\mu + \frac{\sigma^2 t}{n}\right) exp\left\{\mu t + \frac{\sigma^2 t^2}{2n}\right\}$$

e

$$M''_{\bar{X}}(t) = \left(\mu + \frac{\sigma^2 t}{n}\right)^2 exp\left\{\mu t + \frac{\sigma^2 t^2}{2n}\right\} + \left(\frac{\sigma^2}{n}\right) exp\left\{\mu t + \frac{\sigma^2 t^2}{2n}\right\}.$$

Então,

$$E\bar{X} = M'_{\bar{X}}(0) = \mu$$

e

$$E\bar{X}^2 = M''_{\bar{X}}(0) = \mu^2 + \frac{\sigma^2}{n}.$$

E a variância é dada por

$$Var(\bar{X}) = E\bar{X}^2 - (E\bar{X})^2 = \mu^2 + \frac{\sigma^2}{n} - \mu^2 = \frac{\sigma^2}{n}.$$

O que prova o resultado.

2.3 Teorema do Limite Central

O Teorema do Limite Central aproxima grandes amostras da distribuição amostral da média a uma distribuição normal, o que justifica o uso da curva normal em diversas situações.

Antes de demonstrar o Teorema do Limite Central precisamos do seguinte Lema.

Lema 2.11. *Seja $a_1, \ldots, a_n$ uma sequência de números tal que* $\lim_{n\to\infty} \alpha_n = \alpha$*. Então*

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a.$$

Demonstração. Seja

$$L = \lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n$$

Neste caso, temos uma indeterminação do tipo 1∞. Usando logaritmos, vamos transformá- la numa indeterminação do tipo 0/0. Então,

$$\begin{aligned}\ln L &= \ln\left[\lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n\right]\\ &= \lim_{n\to\infty}\left[\ln\left(1+\frac{a_n}{n}\right)^n\right]\\ &= \lim_{n\to\infty} n\ln\left(1+\frac{a_n}{n}\right)\\ &= \lim_{n\to\infty}\frac{\ln(1+a_n/n)}{(1/n)}.\end{aligned}$$

Temos agora, uma indeterminação do tipo 0/0. Aplicando a regra de L'Hospital, obtemos

$$\begin{aligned}\ln L &= \lim_{n\to\infty}\frac{-(a_n/n^2)/(1+a_n/n)}{-(1/n^2)}\\ &= \lim_{n\to\infty}\frac{a_n}{1+a_n/n}\\ &= a.\end{aligned}$$

Dessa forma, $L = e^a$ e, portanto,

$$\lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n = e^a.$$

Teorema 2.12 (Teorema do Limite Central)**.** *Sejam $X_1, \ldots, X_n$ variáveis aleatórias i.i.d., tais que as f.g.m., $M_{X_i}(t)$, existam para $|t| < h$, $h > 0$. Sejam $EX_i = \mu$ e $Var(X_i) = \sigma^2 > 0$, ambos finitos. Defina $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Então, para qualquer $x \in \mathbb{R}$,*

$$\lim_{n\to\infty} P\left(\frac{\sqrt{n}\,(\bar{X}_n-\mu)}{\sigma} < x\right) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{\frac{-y^2}{2}}\,dy.$$

Demonstração. Mostraremos que, para $|t| < h$, a f.g.m. de $\sqrt{n}\,(\bar{X}_n - \mu)/\sigma$ converge para a f.g.m. de $X \sim N(0, 1)$.

Tome $(X_i - \mu)/\sigma = X_i^*$. Temos,

$$\frac{\sqrt{n}\,(\bar{X}_n-\mu)}{\sigma} = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right)}{\sigma} = \frac{\left(\sum_{i=1}^n X_i - n\mu\right)}{\sqrt{n}\sigma} = \frac{\sum_{i=1}^n (X_i-\mu)}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i^*.$$

Então, a f.g.m. de $\sqrt{n}\,(\bar{X}_n - \mu)/\sigma$ é igual a f.g.m. de $\sum_{i=1}^n X_i^*/\sqrt{n}$.

Basta mostrar que, para $|t| < h$, a f.g.m. de $\sum_{i=1}^n X_i^*/\sqrt{n}$ converge para a f.g.m. de $X \sim N(0, 1)$.

Pelo Teorema 1.45 e pela Equação (1.63), temos

$$M_{\sum_{i=1}^{n} X_i^*/\sqrt{n}}(t) = M_{\sum_{i=1}^{n} X_i^*}\left(\frac{t}{\sqrt{n}}\right) = \left[M_{X^*}\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

Expandindo $M_{X^*}(t/\sqrt{n})$, em uma série de Taylor em torno de 0, obtemos

$$M_{X^*}\left(\frac{t}{\sqrt{n}}\right) = \sum_{k=0}^{\infty} M_{X^*}^{(k)}(0)\frac{\left(t/\sqrt{n}\right)^k}{k!}.$$

Uma vez que, por construção, a média e a variância são, respectivamente 0 e 1, obtemos,

$$M_{X^*} = 1,$$
$$M'_{X^*} = 0,$$

e

$$M''_{X^*} = 1,$$

Então,

$$M_{X^*}\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{\left(t/\sqrt{n}\right)^2}{2!} + R_{X^*}\left(\frac{t}{\sqrt{n}}\right),$$

onde definimos $R_{X^*}(t/\sqrt{n})$ como sendo o termo restante da aproximação do polinômio de Taylor.

Como $M_{X^*}^{(k)}(0) = \frac{d^k}{d_x^k} M_{X^*}(t)|_{t=0}$ existe, e admitindo a utilização do Teorema de Taylor, $\lim_{x \to c}[f(x) - T_r(x)]/(x-c)^r = 0$, então, para qualquer $t \neq 0$ fixo, obtemos o seguinte resultado:

$$\lim_{n\to\infty} \frac{R_{X^*}(t/\sqrt{n})}{\left(t/\sqrt{n}\right)^2} = 0.$$

E, em particular, para $t = 1$,

$$\lim_{n\to\infty} \frac{R_{X^*}(t/\sqrt{n})}{\left(1/\sqrt{n}\right)^2} = \lim_{n\to\infty} nR_{X^*}\left(\frac{t}{\sqrt{n}}\right) = 0.$$

Para $t = 0$, temos $R_{X^*}(0/\sqrt{n}) = 0$.

Logo, para todo $t \in$ R, temos

$$\lim_{n\to\infty}\left[M_{X^*}\left(\frac{t}{\sqrt{n}}\right)\right]^n = \lim_{n\to\infty}\left[1+\frac{(t/\sqrt{n})^2}{2!}+R_{X^*}\left(\frac{t}{\sqrt{n}}\right)\right]^n$$
$$= \lim_{n\to\infty}\left[1+\frac{1}{n}\left(\frac{t^2}{2}+nR_{X^*}\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n.$$

Fazendo $[(t^2/2) + nR_{X^*}(t/\sqrt{n})] = \alpha_n$ no Lema 2.11, obtemos

$$\lim_{n\to\infty}\left[M_{X^*}\left(\frac{t}{\sqrt{n}}\right)\right]^n = e^{\frac{t^2}{2}}$$

que é a f.g.m. de $X \sim N(0, 1)$, o que prova o resultado.

2.4 Aproximação normal para a distribuição binomial

O Teorema a seguir nos diz que quando n é grande, $X \sim Binomial(n, p)$ tem aproximadamente a mesma distribuição que $X \sim N(np, np(1 - p))$.

Teorema 2.13. *(Teorema Limite de DeMoivre e Laplace) Seja S_n o número de sucessos em n tentativas individuais, cada uma com probabilidade p de sucesso, então para $a < b$*

$$\lim_{n\to\infty} P\left\{a \leqslant \frac{S_n - np}{\sqrt{np(1-p)}} \leqslant b\right\} = \Phi(b) - \Phi(a).$$

Demonstração. Este Teorema é um caso particular do Teorema do Limite Central.

$$\lim_{n\to\infty} P\left\{a \leqslant \frac{S_n - np}{\sqrt{np(1-p)}} \leqslant b\right\} = \int_a^b \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz = \Phi(b) - \Phi(a).$$
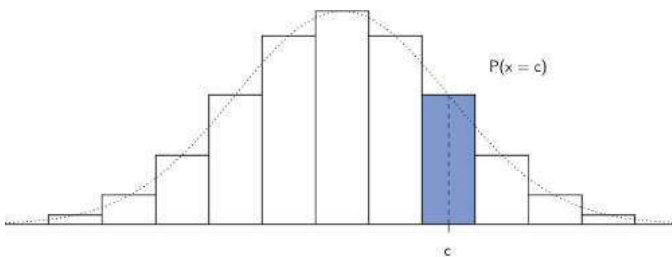
Figura 2: Probabilidade binomial exata

Fonte: Arquivo próprio do autor, baseado em [8].

Quando aproximamos uma distribuição normal contínua para uma distribuição binomial, que é discreta, devemos fazer uma correção pela continuidade, que consiste em mover 0, 5 unidade tanto à esquerda quanto à direita, a fim de incluir todos os valores possíveis de x no intervalo.
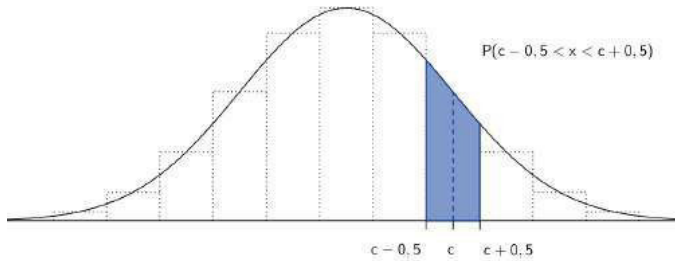
Figura 3: Aproximação normal da binomial

Fonte:Arquivo próprio do autor, baseado em [8].

Assim, escrevemos a distribuição binomial $P(X = i)$ como $P(i - 0,5 < X < i + 0,5)$ ao aproximar da normal.

Podemos fazer tal aproximação quando $np \geq 5$ e $n(1 - p) \geq 5$.

2.5 Distribuição amostral de uma proporção

Considerando uma população tal que p é a proporção de elementos com uma determinada característica, definimos a variável aleatória X a seguir,

$$X = \begin{cases} 1, \text{ se o elemento for portador da característica} \\ 0, \text{ se o elemento não for portador da característica.} \end{cases}$$

Retirada uma amostra aleatória, com reposição, e denotando por S_n o total de elementos portadores da característica, temos,

$$S_n = \sum_{i=1}^{n} X_i, \text{ com } i = 1, \cdots + n.$$

onde cada X_i, tem distribuição de Bernoulli, com média $\mu = p$ e variância $\sigma^2 = p(1 - p)$, independentes duas a duas, então,

$$S_n = n\bar{X}.$$

Pelo Teorema do Limite Central, temos,

$$\bar{X} \sim N\left(p, \frac{p(1-p)}{n}\right). \tag{2.7}$$

Logo,

$$S_n \sim N(np, np(1-p)). \tag{2.8}$$

Denotando por $\hat{p}$ a proporção de elementos portadores da característica na amostra, obtemos,

$$\hat{p} = \frac{S_n}{n}. \tag{2.9}$$

Dessa forma,

$$P(S_n = k) = P(S_n/n = k/n) = P(\hat{p} = k/n).$$

E como, $\bar{X} = \hat{p}$, temos,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right). \tag{2.10}$$

2.6 Distribuição da variância da amostral

A variância amostral, definida em 2.5, é uma estatística. Podemos calcular *ES*2. Para calcular ES^2 precisamos do Teorema a seguir.

Teorema 2.14. *Sejam* $x_1, \ldots, x_n$ *números reais e* $\bar{x} = \frac{x_1+\cdots+x_n}{n}$, *Então,*

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2, \text{ com } i = 1, \ldots, n. \tag{2.11}$$

Demonstração. Tome $\sum_{i=1}^{n} x_i^2$ e adicione e subtraia $\bar{x}$. Obtendo, dessa forma,

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n}(x_i - \bar{x} + \bar{x})^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 + 2\sum_{i=1}^{n}(x_i - \bar{x})\bar{x} + \sum_{i=1}^{n}\bar{x}^2.$$

Observando que $\sum_{i=1}^{n} (x_i - \bar{x}) = 0$ e que $\sum_{i=1}^{n} \bar{x}^2 = n\bar{x}^2$, obtemos (2.11)

Teorema 2.15. *Sejam* $X_1, \ldots, X_n$ *uma amostra aleatória de uma população com EX* = μ *e* $Var(X) = \sigma^2 < \infty$. *Então, a variância amostral é dada por*

$$ES^2 = \sigma. \tag{2.12}$$

Demonstração. Pela Definição 2.5, e do Teorema 2.14, temos

$$ES^2 = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = E\left(\frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]\right) = \frac{1}{n-1}\left(nEX^2 - nE\bar{X}^2\right).$$

Da Equação (1.18), obtemos

$$E(X^2) = Var(X) - (EX)^2.$$

Logo, de (2.3) e (2.4),

$$ES^2 = \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right] = \sigma^2.$$

3 | ESTIMADORES

Seja $X_1, \ldots, X_n$ uma amostra de uma população com f.p. ou f.d.p. $f(x|\theta)$, temos a seguinte definição.

Definição 2.16. O estimador pontual T de um parâmetro θ é qualquer função $g(X_1, \ldots, X_n)$ de uma amostra, ou seja, $T = g(X_1, \ldots, X_n)$.

Um estimador é uma estatística associada a um parâmetro.

Uma estimativa é um valor observado de um estimador.

Definição 2.17. Uma estatística $T = g(X_1, \ldots, X_n)$ é suficiente para θ, quando a distri- buição condicional de $X_1, \ldots, X_n$, dado T não depender de θ.

3.1 Estimadores de Momentos

Seja $X_1, \ldots, X_n$ uma amostra de uma população X com f.p. ou f.d.p. $f(x|\theta_1, \ldots, \theta_k)$, dependendo de k parâmetros, temos as seguintes definições:

Definição 2.18. Definimos o k-ésimo momento de X por

$$\mu' k = EX^k, \text{ com } k = 1, \ldots, n.$$

Definição 2.19. Definimos o k-ésimo momento amostral por

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \text{ com } k = 1, \ldots, n.$$

Definição 2.20. Dizemos que $\hat{\theta}1, \ldots, \hat{\theta}k$ são estimadores obtidos pelo método dos mo- mentos se eles forem solução do seguinte sistema de equações.

$$\begin{aligned} m_1 &= \mu 1' (\theta_1, \ldots, \theta_k), \\ m_2 &= \mu 2' (\theta_1, \ldots, \theta_k), \\ &\vdots \\ m_k &= \mu k' (\theta_1, \ldots, \theta_k). \end{aligned}$$

Exemplo 2.1. Método dos momentos da normal.

Seja $X \sim N(\mu, \sigma^2)$, então $\theta_1 = \mu$ e $\theta_2 = \sigma^2$. Temos, $\mu 1' = EX = \mu$, $\mu_2' = EX^2 = \sigma^2 + \mu^2$, $m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$ e $m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$. Então,

$$\bar{X} = \mu$$

e

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \sigma^2 + \mu^2.$$

Os estimadores obtidos pelo método dos momentos serão

$$\hat{\mu}_M = \bar{X}$$

e

$$\hat{\sigma}^2_M = m_2 - m_1^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 .$$

3.2 Estimadores de máxima verossimilhança

Seja $X_1, \ldots, X_n$ uma amostra i.i.d. de uma população X com f.p. ou f.d.p. $f(x|\theta_1, \ldots, \theta_k)$, temos a seguinte definição.

Definição 2.21. A função de verossimilhança é definida por

$$L(\theta|x) = L(\theta_1, \ldots, \theta_k|x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\theta_1, \ldots, \theta_k).$$

Exemplo 2.2. Exemplo extraído de [17].

Verossimilhança da média de uma distribuição normal com variância conhecida.

Seja $X \sim N(\mu, \sigma^2)$, com σ conhecido e μ desconhecido, a função de verossimilhança de μ é

$$L(\mu) = f(x|\mu) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\mu}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{\sigma\sqrt{2\mu}}\right)^n e^{-\sum\frac{(x_i-\mu)^2}{2\sigma^2}}.$$

Essa função tem valor máximo quando

$$g(\mu) = \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

é mínimo. Derivando e fazendo $g'(\mu) = 0$, obtemos

$$g'(\mu) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i - \mu) = 0.$$

De onde

$$\sum_{i=1}^{n} x_i - n\mu = 0.$$

Consequentemente,

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}_n.$$

é o valor crítico de μ. Desde que $g''(\mu) = n/\sigma^2 > 0$, a função g tem um mínimo e a função L, um máximo em $\hat{\mu} = \bar{x}_n$.

3.3 Erro quadrático médio

Definição 2.22. O erro quadrático médio (EQM) de um estimador T de um parâmetro θ é a função de θ definida por $E_\theta(T - \theta)^2$.

Definimos a tendência de um estimador T do seguinte modo.

Definição 2.23. O viés de um estimador T de um parâmetro θ é $viés_\theta\ T = E_\theta\ T - \theta$, ou seja, é a diferença entre o valor esperado de T e θ. Um estimador T é não-viesado se $E_\theta\ T = \theta$, e seu viés é igual a zero.

Da Definição 2.22,e adicionando e subtraindo $E_\theta\ T$, temos

$$\begin{aligned} E_\theta(T - \theta)^2 &= E_\theta(T - E_\theta\ T + E_\theta\ T - \theta)^2 \\ &= E_\theta(T - E_\theta\ T)^2 + 2E_\theta[(T - E_\theta\ T)(E_\theta\ T - \theta)] + E_\theta(E_\theta\ T - \theta)^2. \end{aligned}$$

Como $E_\theta\ T - \theta$ é uma constante e $E_\theta(T - E_\theta\ T) = 0$, obtemos

$$E_\theta(T - \theta)^2 = E_\theta(T - E_\theta\ T)^2 + E_\theta(E_\theta\ T - \theta)^2.$$

Das Definições 1.15 e 2.23, segue

$$E_\theta(T - \theta)^2 = Var_\theta(T) - vie{'}s_\theta\ T.$$

Se um estimador T é não-viesado, seu EQM é igual a sua variância, ou seja,

$$E_\theta(T - \theta)^2 = Var_\theta(T).$$

Exemplo 2.3. Seja X uma população de tamanho N com variância populacional $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$, com média populacional $\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$. Extraída uma amostra aleatória simples de tamanho n, temos que $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, é um possível estimador para σ^2.

Queremos mostrar que este estimador é viesado, para tal, procedemos de maneira análoga a utilizada no Teorema 2.15.

Do Teorema 2.14, temos

$E(\ ^2) = E\quad \underline{1}\ n\ (X$

$$E(\sigma^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = E\left(\frac{1}{n}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]\right) = \frac{1}{n}\left(nEX^2 - nE\bar{X}^2\right).$$

Então,

$$E(\sigma^2) = \frac{1}{n}\left[n(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right] = \frac{n-1}{n}\sigma^2.$$

Portanto, σ^2 é um estimador viesado.

Como $E\bar{X} = \mu$ e $ES^2 = \sigma^2$, temos que $\bar{X}$ e S^2 são estimadores não-viesados para μ e σ^2, respectivamente. Por essa razão, escrevemos $n - 1$ no denominador da variância amostral.

Com o objetivo de se encontrar o melhor estimador não-viesado, temos a seguinte definição.

Definição 2.24. Dados dois estimadores não-viesados, T e T^*, de um mesmo parâmetro θ, dizemos que T é um melhor estimador, ou um estimados mais eficiente do que T^* se $Var(T) < Var(T^*)$.

4 | INTERVALOS DE CONFIANÇA

Os intervalos de confiança baseiam-se na distribuição amostral do estimador pontual. Os resultados a seguir são frequentemente utilizados em sua construção, bem como nos testes de hipóteses para populações normais.

Teorema 2.25. *Sejam $X_1, \ldots, X_n$ uma amostra aleatória de tamanho n da distribuição* $N \sim (\mu, \sigma^2)$, *e* $\bar{X} = \sum_{i=1}^n X_i/n$ *e* $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$. *Então,*

(i) $\bar{X}$ *e* S^2 ão *variáveis aleatórias independentes;*

(ii) $(n - 1)S^2/\sigma^2$ *tem uma distribuição qui-quadrada com* $n - 1$ *graus de liberdade;*

(iii) $\sqrt{n}(\bar{X} - \mu)/S$ *tem uma distribuição t de Student com* $n - 1$ *graus de liberdade.*

Demonstração. A demonstração foi extraída de [2].

(i) Temos que

$$\bar{X} \sim N(\mu, \sigma^2/n),$$

enquanto que $X_i - \bar{X} \sim N\left(0, \sigma^2 \frac{(n-1)}{n}\right)$. Por outro lado, a função geradora de momentos de $Y_1 = \bar{X}$ e $Y_2 = X_i - \bar{X}$ é dada por

$$\begin{aligned} M_{Y_1,Y_2}(s_1,s_2) &= E\left[e^{s_1\bar{X}+s_2(X_i-\bar{X})}\right] = E\left[e^{s_2X_i+\bar{X}(s_1-s_2)}\right] \\ &= E\left[e^{\left(s_2+\frac{(s_1-s_2)}{n}\right)X_i+\frac{(s_1-s_2)}{n}\sum_{j\neq i}^n X_j}\right] \\ &= E\left[e^{\left(s_2+\frac{(s_1-s_2)}{n}\right)X_i}\right] E\left[e^{\frac{(s_1-s_2)}{n}\sum_{j\neq i}^n X_j}\right] \end{aligned}$$

Como $X_i \sim N(\mu, \sigma^2)$ e $\sum_{j\neq i}^n X_j \sim N((n-1)\mu; (n-1)\sigma^2)$, temos que

$$
\begin{aligned}
M_{Y_1,Y_2}(s_1,s_2) &= exp\left\{\mu\left(s_2+\frac{(s_1-s_2)}{n}\right)+\sigma^2\left(s_2+\frac{(s_1-s_2)}{n}\right)^2\right\} \\
&\cdot \quad exp\left\{\frac{(n-1)}{n}(s_1-s_2)\mu+\frac{1}{2}\left\{\left(\frac{s_1-s_2}{n}\right)^2(n-1)\sigma^2\right\}\right\} \\
&= exp\left\{\mu s_1+\frac{s_1^2\sigma^2}{2n}\right\} exp\left\{\frac{s_2^2(n-1)\sigma^2}{2n}\right\}
\end{aligned}
$$

que é o produto das funções geradoras de momentos das distribuições de $\bar{X}$ e $X_i - \bar{X}$. Portanto temos que $X_i - \bar{X}$ e $\bar{X}$ são independentes, pois a função geradora da distribuição conjunta é o produto das funções geradoras de momentos das distribuições marginais. Como $\sum_{i=1}^{n}(X_i - \bar{X})^2$ é função de $X_i - \bar{X}$ que é independente de $\bar{X}$, temos que S^2 é independente de $\bar{X}$.

(ii) Temos que

$$
\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}=\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\sigma^2}+n\frac{(\bar{X}-\mu)^2}{\sigma^2}. \qquad (2.13)
$$

Como $(X_i - \mu)/\sigma \sim N(0, 1)$, temos que $(X_i - \mu)^2/\sigma^2 \sim \chi_1^2$, $i = 1, \ldots, n$, de modo que

$$
Y_1=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\sim\chi_n^2.
$$

Também $n(X_i - \mu)^2/\sigma^2 \sim \chi_1^2$. Como a função geradora de momentos da distribuição qui-quadrado com g graus de liberdade é dada por

$$
M_g(s) = (1 - 2s)^{-g/2},
$$

temos que as funções geradoras das distribuições qui-quadrado com $g = 1$ e $g = n$ graus de liberdade são dadas respectivamente por

$$
M_1(s) = (1 - 2s)^{-1/2} \text{ e } M_n(s) = (1 - 2s)^{-n/2}, \qquad (2.14)
$$

Além disso, como $\bar{X}$ e S^2 são independentes, temos que os dois termos do lado direito de (2.13) que dentamos por Y_2 e Y_3, respectivamente, são independentes, de modo que

$$
M_{Y_1}(s) = M_{Y_2}(s)M_{Y_3}(s),
$$

ou seja, de (2.14) segue que

$$
M_{Y_2}(s)=\frac{M_{Y_1}(s)}{M_{Y_3}(s)}=(1-2s)^{-(n-1)/n},
$$

logo o distribuição de $Y_2 = (n-1)S^2/\sigma^2$ é qui-quadrada com $n-1$ graus de liberdade.

(iii) Note que podemos escrever

$$\sqrt{n}\frac{(\bar{X}-\mu)}{S} = \frac{\sqrt{n}\frac{(\bar{X}-\mu)}{\sigma}}{\sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}}, \tag{2.15}$$

que corresponde ao quociente entre duas variáveis aleatórias independentes em que o numerador é uma variável aleatória com distribuição $N(0, 1)$ e o denominador é a raiz quadrada de uma variável aleatória com distribuição qui-quadrada com $n-1$ graus de liberdade (veja(ii)) dividido pelo pelo número de graus de liberdade, de modo que a variável (2.15) tem distribuição t de Student com $n-1$ graus de liberdade.

Definição 2.26. Uma variável aleatória T tem distribuição t de Student com n graus de liberdade, $T \sim t_n$ se sua f.d.p. for dada por

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{(n\pi)^{1/2}} \frac{1}{(1+t^2/n)^{(n+1)/2}}, \quad -\infty < t < \infty. \tag{2.16}$$

4.1 Intervalo de confiança para a média, com variância conhecida

Ao obtermos uma amostra, sua média μ pode não ser exatamente a média populacional. Podemos estimar que μ pertence a um intervalo, desta forma, temos uma estimativa intervalar.

Definição 2.27. Uma estimativa intervalar é um intervalo utilizado para estimar um parâmetro populacional.

Quando tomamos $\bar{x}$ como uma estimativa de μ, precisamos determinar o quanto estamos confiantes de que $\bar{x}$ pertença ao intervalo estimado.

Definimos, a seguir, o nível de confiança para um parâmetro populacional.

Definição 2.28. O nível de confiança $1-\alpha$ é a probabilidade de que o parâmetro populacional pertença ao intervalo estimado.
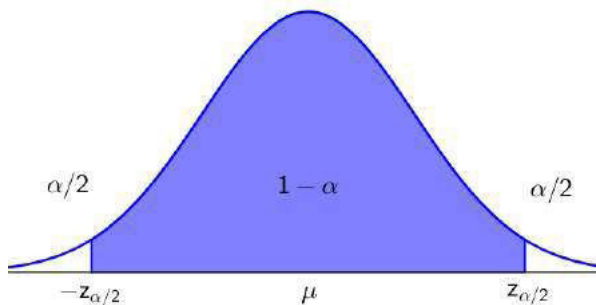
Figura 4: Distribuição normal com nível de confiança $1-\alpha$

Fonte: Arquivo próprio do autor, baseado em [10].

O nível de confiança é a área entre os valores críticos $-z_{\alpha/2}$ e $z_{\alpha/2}$ na Figura 4, obtidos na Tabela 2.

A margem de erro E é a maior diferença possível entre o valor estimado e o verdadeiro valor do parâmetro, e é dada por:

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

A distribuição amostral da média se aproxima de uma distribuição normal com média μ e variância σ^2/n, de modo que

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

é o valor de uma variável aleatória que tem aproximadamente uma distribuição normal padrão, com probabilidade $1 - \alpha$ de que

$$-z_{\alpha/2} < z < z_{\alpha/2}.$$

Substituindo a Equação anterior, temos

$$-z_{\alpha/2} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}.$$

Multiplicando cada termo por $\sigma/\sqrt{n}$, subtraindo $\bar{x}$ e multiplicando o resultado ob- tido por -1, obtemos

$$\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} > \mu > \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Logo, podemos escrever:

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Este intervalo é conhecido como intervalo de confiança para a média, com variância conhecida. E, neste caso,

$$P\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Para uma mesma amostra, à medida que o nível de confiança aumenta, o inter- valo de confiança fica mais largo e a precisão da estimativa decresce. Aumentando o tamanho da amostra aumentamos a precisão da estimativa. Logo, dados o nível de confiança $1 - \alpha$ e uma margem de erro E pré-determinada, o tamanho mínimo da amostra n para estimar a média populacional, com σ^2 conhecido é:

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 .$$

4.2 Intervalo de confiança para a média, com variância desconhecida

Geralmente, o desvio padrão populacional é desconhecido. Sendo a população normalmente distribuída, ou aproximadamente normalmente distribuída, pelo Teorema 2.25 (iii), temos que

$$t_{n-1} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

é o valor de uma variável aleatória com distribuição *t* de Student com $n - 1$ graus de liberdade.
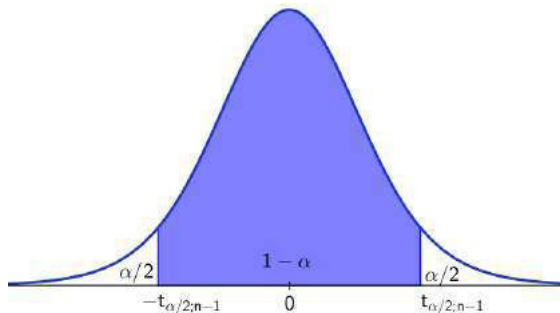
Figura 5: Distribuição *t* de Student com nível de confiança $1 - \alpha$

Fonte: Arquivo próprio do autor, baseado em [10].

O nível de confiança é a área entre os valores críticos $-t_{\alpha/2;n-1}$ e $t_{\alpha/2;n-1}$ na Figura 5, obtidos na Tabela 3.

Temos, com probabilidade $1 - \alpha$ que

$$-t_{\alpha/2;n-1} < t < t_{\alpha/2;n-1}.$$

Substituindo a Equação anterior, temos

$$-t_{\alpha/2;n-1} < \frac{\bar{x} - \mu}{s/\sqrt{n}} < t_{\alpha/2;n-1}.$$

Multiplicando cada termo por $s/\sqrt{n}$, subtraindo $\bar{x}$ e multiplicando o resultado obtido por -1, obtemos

$$\bar{x} + t_{\alpha/2;n-1}\frac{s}{\sqrt{n}} > \mu > \bar{x} - t_{\alpha/2;n-1}\frac{s}{\sqrt{n}}.$$

Logo, podemos escrever:

$$\bar{x} - t_{\alpha/2;n-1}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2;n-1}\frac{s}{\sqrt{n}}.$$

Este intervalo é conhecido como intervalo de confiança para a média, com variância desconhecida. E, neste caso,

$$P\left(\bar{x} - t_{\alpha/2;n-1}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2;n-1}\frac{s}{\sqrt{n}}\right) = 1 - \alpha.$$

4.3 Intervalo de confiança para a variância e o desvio padrão

Pode haver o interesse em fazer estimativas sobre outros parâmetros populacionais, como a variância e o desvio padrão. Podemos definir as estimativas pontuais para estes parâmetros.

Definição 2.29. As estimativas pontuais para σ^2 e σ são, respectivamente, s^2 e s. Te- mos, ainda, que $s2$ é a melhor estimativa não tendenciosa para σ^2.

Para uma população que tenha aproximadamente uma distribuição normal, temos, pelo Teorema 2.25 (ii), que

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma^2}$$

é o valor de uma variável aleatória com distribuição qui-quadrado com $n - 1$ graus de liberdade.
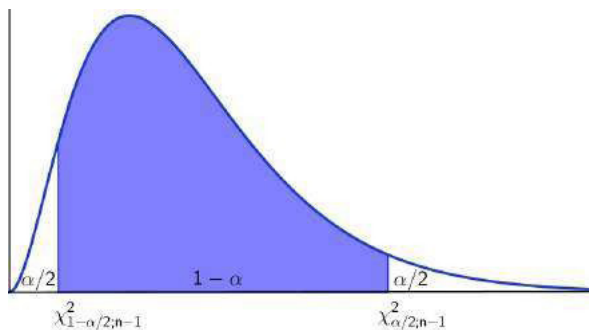
Figura 6: Distribuição qui-quadrado com nível de confiança 1 – α

Fonte: Arquivo próprio do autor, baseado em [10].

O nível de confiança é a área entre os valores críticos $\chi^2_{1-\alpha/2;n-1}$ e $\chi^2_{\alpha/2;n-1}$ na Figura 6, obtidos na Tabela 4.

A distinção entre $\chi^2_{1-\alpha/2;n-1}$ e $\chi^2_{\alpha/2;n-1}$ se faz necessária pois a curva da distribuição qui-quadrado não é simétrica.

Temos, com probabilidade 1 – α que

$$\chi^2_{1-\alpha/2;n-1} < \chi^2 < \chi^2_{\alpha/2;n-1}$$

Substituindo a Equação anterior, temos

$$\chi^2_{1-\alpha/2;n-1} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2;n-1},$$

que podemos reescrever como

$$\frac{1}{\chi^2_{1-\alpha/2;n-1}} > \frac{\sigma^2}{(n-1)s^2} > \frac{1}{\chi^2_{\alpha/2;n-1}}.$$

Multiplicando cada termo por $(n-1)s^2$, obtemos

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2;n-1}} > \sigma^2 > \frac{(n-1)s^2}{\chi^2_{\alpha/2;n-1}}.$$

Logo, podemos escrever:

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2;n-1}}$$

Este intervalo é conhecido como intervalo de confiança para a variância. E, neste caso,

$$P\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2;n-1}}\right) = 1-\alpha.$$

Podemos obter um intervalo de confiança para o desvio padrão:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2;n-1}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2;n-1}}}$$

Dessa forma,

$$P\left(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2;n-1}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2;n-1}}}\right) = 1-\alpha.$$

4.4 Intervalo de confiança para uma proporção

Da Equação (2.9), temos que

$$\hat{p} = \frac{S_n}{n}$$

é uma estimativa pontual para p e a estatística

$$z = \frac{x - np}{\sqrt{np(1-p)}}.$$

é o valor de uma variável aleatória que tem aproximadamente uma distribuição normal padrão, com probabilidade 1 – α de que

$$-z_{\alpha/2} < z < z_{\alpha/2}.$$

Substituindo a Equação anterior, temos

$$-z_{\alpha/2} < \frac{x - np}{\sqrt{np(1-p)}} < z_{\alpha/2}.$$

Multiplicando cada termo por $\sqrt{np(1-p)}$, subtraindo x e multiplicando o resultado obtido por –1, obtemos

$$x + z_{\alpha/2}\sqrt{np(1-p)} > np > x - z_{\alpha/2}\sqrt{np(1-p)}.$$

Dividindo cada termo por n e rearranjando, obtemos,

$$\frac{x}{n} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < p < \frac{x}{n} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}.$$

Sendo p um parâmetro desconhecido e de (2.7) segue que

$$\sqrt{\frac{p(1-p)}{n}}$$

é o desvio padrão da distribuição amostral de uma proporção.

Fazendo $S_n = x$ em (2.9) e substituindo

$$\hat{p} = \frac{x}{n} \text{ para } p \text{ em } \sqrt{\frac{p(1-p)}{n}}.$$

obtemos um intervalo de confiança para uma proporção,

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Logo,

$$P\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1 - \alpha.$$

5 I TESTES DE HIPÓTESE

Para iniciar esta seção definimos hipótese estatística.

Definição 2.30. Uma hipótese estatística é uma afirmação sobre um parâmetro populacional.

Em um teste de hipótese devemos decidir, a partir de uma amostra populacional, entre duas hipóteses complementares, qual é verdadeira.

Definição 2.31. As duas hipóteses complementares são chamadas *hipótese nula*, H_0 e *hipótese alternativa*, H_1.

Seja θ um parâmetro populacional e afirmamos que $\theta = \theta_0$, onde θ_0 é a hipótese nula. Temos, então,

$$H_0 : \theta = \theta_0.$$

De forma mais geral, a hipótese alternativa seria

$$H_1 : \theta \neq \theta_0.$$

É possível termos hipóteses alternativas na forma

$$H_1 : \theta < \theta_0 \text{ ou } H_1 : \theta > \theta_0.$$

Devemos, então, decidir se aceitamos H_0 como verdadeira ou se a rejeitamos, aceitando H_1 como verdadeira.

Definição 2.32. Para testar uma hipótese devemos saber para quais valores amostrais aceitamos H_0 como verdadeira e para quais valores amostrais rejeitamos H_0 e aceita- mos H_1 como verdadeira.

Definimos, a seguir, a região crítica.

Definição 2.33. A *região crítica* ou *região de rejeição* é o subconjunto do espaço amos- tral para o qual rejeitamos H_0, seu complemento é chamado de *região de aceitação*.

Ao testar uma hipótese dois tipos de erros podem ocorrer, o erro do tipo I e o erro do tipo II.

O erro do tipo I ocorre quando rejeitamos H_0 sendo H_0 verdadeira e o erro do tipo II ocorre quando não rejeitamos H_0 sendo H_0 falsa.

A Tabela 1 descreve esta situação:

Decisão	H_0 é verdadeira	H_0 é falsa
Aceitar H_0 Rejeitar H_0	Decisão correta Erro do tipo II	Erro do tipo I Decisão correta

Tabela 1: Tipos de erros em testes de hipóteses

Fonte: [2].

A probabilidade máxima permissível para cometer um erro do tipo I é denominado *nível de significância*.

As probabilidades de cometer os erros do tipo I e do tipo II são denotadas α e β, respectivamente.

$$\alpha = P(\text{erro do tipo I}).$$
$$\beta = P(\text{erro do tipo II}).$$

Seja RC a região crítica para um teste de hipótese. O teste resultará em um erro, para $\theta = \theta_0$, se $x \in RC$, dessa forma, a probabilidade de um erro do tipo I é $P_\theta(X \in RC)$. E a probabilidade de um erro do tipo II é $1 - P_\theta(X \in RC)$. Temos

$$P_\theta\,(X \in RC) = \begin{cases} \text{probabilidade de erro do tipo I, se } \theta = \theta_0; \\ \text{um menos a probabilidade de erro do tipo II, caso contrário.} \end{cases}$$

O que nos leva à seguinte definição.

Definição 2.34. A *função poder* de um teste de hipótese com região crítica RC é a função de θ definida por $\beta(\theta) = P_\theta(X \in RC)$.

5.1 Teste de hipótese para a média com variância conhecida

O teste de hipótese para a média com variância conhecida é o teste z para a média μ.

Sejam $\bar{x}$ a média amostral, μ_0 o valor da média sob a hipótese nula e σ^2/n a variância. A estatística do teste padronizado é z.

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Consideramos $H_0 : \mu = \mu_0$.
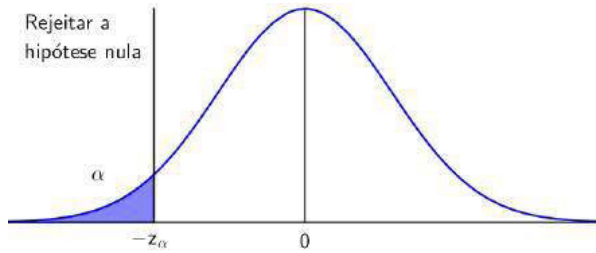
Figura 7: Teste z unicaudal à esquerda, $H_1 : \mu < \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].
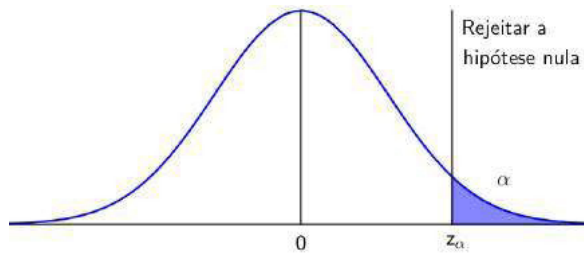
Figura 8: Teste z unicaudal à direita, $H_1 : \mu > \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].
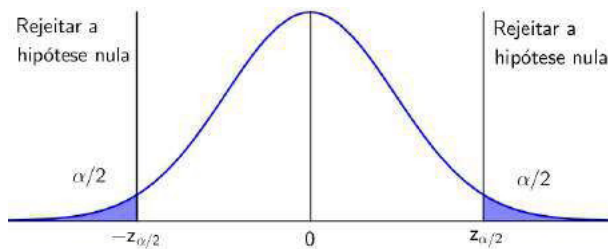
Figura 9: Teste z bicaudal, $H_1 : \mu /= \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].

5.2 Teste de hipótese para a média com variância desconhecida

O teste de hipótese para a média com variância desconhecida com a população normalmente distribuída, ou aproximadamente normalmente distribuída, é o teste t para a média μ.

Pelo Teorema 2.25 (iii) , temos que

$$t_{n-1} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

é o valor de uma variável aleatória com distribuição t de Student com $n - 1$ graus

de liberdade.

Do mesmo modo como na seção anterior, consideramos $H_0 : \mu = \mu_0$.
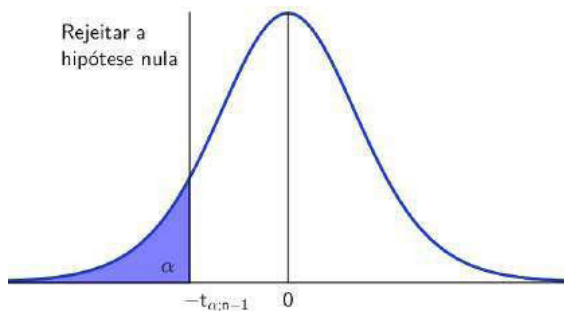
Figura 10: Teste *t* unicaudal à esquerda, $H_1 : \mu < \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].
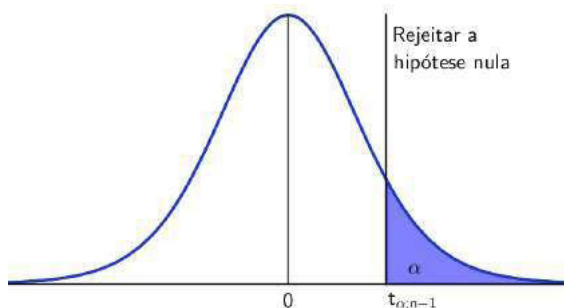
Figura 11: Teste *t* unicaudal à direita, $H_1 : \mu > \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].
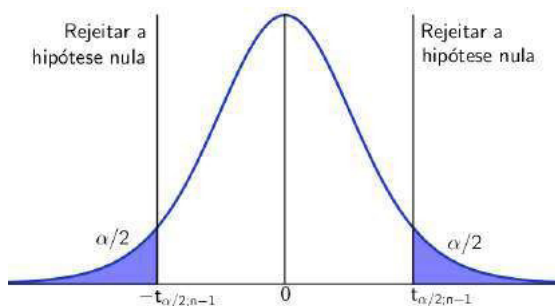
Figura 12: Teste *t* bicaudal, $H_1 : \mu /= \mu_0$

Fonte: Arquivo próprio do autor baseado em [10].

5.3 Teste de hipótese para a variância

O teste de hipótese para a variância, com a população normalmente distribuída, ou aproximadamente normalmente distribuída, é o teste χ^2.

Pelo Teorema 2.25 (ii) , temos que

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma^2}$$

é o valor de uma variável aleatória com distribuição qui-quadrado com $n - 1$ graus de liberdade.

Consideramos $H_0 : \sigma = \sigma_0$.
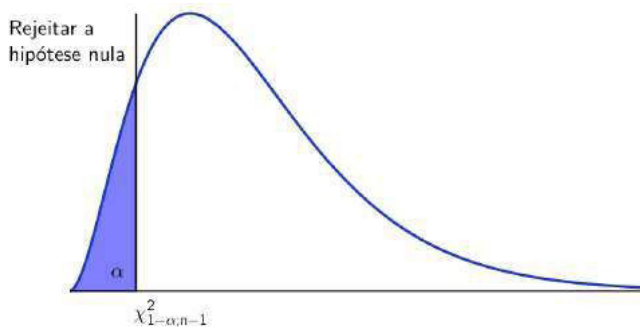
Figura 13: Teste χ^2 unicaudal à esquerda, $H_1 : \sigma < \sigma_0$

Fonte: Arquivo próprio do autor baseado em [10].
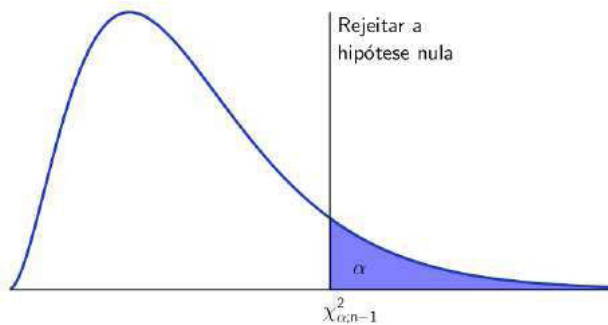
Figura 14: Teste χ^2 unicaudal à direita, $H_1 : \sigma > \sigma_0$

Fonte: Arquivo próprio do autor baseado em [10].
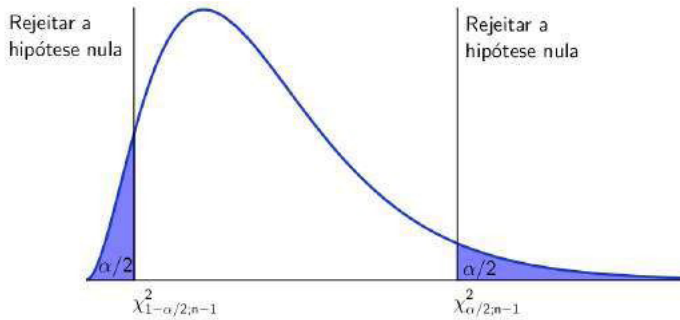
Figura 15: Teste χ^2 bicaudal, $H_1 : \sigma \,/\!= \sigma_0$

Fonte: Arquivo próprio do autor baseado em [10].

5.4 Teste de hipótese para uma proporção

Uma proporção p com $np \geq 5$ e $n(1 - p) \geq 5$ tem uma distribuição binomial que pode ser aproximada a uma distribuição normal. Neste caso, o teste estatístico a ser utilizado é o teste z.

$$z = \frac{\hat{p} - p_0}{\sqrt{p(1-p)/n}}.$$

Onde p_0 é o valor da proporção sob a hipótese nula, $H_0 : p = p_0$.

A rejeição da hipótese nula ocorre como nas Figuras 7, 8 e 9.

PROPOSTA DE ATIVIDADES

Este Capítulo tem como finalidade apresentar uma proposta de atividades sobre inferência no ensino médio, especificamente aos alunos do 3º ano.

Sendo que ao término de cada atividade proposta, os alunos tenham condições de realizarem inferências sobre uma determinada amostra.

Serão desenvolvidas diversas atividades, todas com objetivos específicos.

A primeira atividade, por exemplo, consiste em motivar os jovens à introdução de conceitos mais complexos.

Para a realização das atividades, a metodologia aplicada será através de questionários, como também de diversos cálculos propostos.

Assim, o exercício 1 tem como objetivo primordial relembrar o cálculo para uma distribuição binomial.

No entanto, para a efetiva realização da atividade 1, o aluno deverá possuir prévio conhecimento de como proceder no cálculo de probabilidades binomiais, e se necessário, retomar explicações sobre o assunto.

No exercício 2, será apresentada a curva da distribuição normal, bem como a Tabela 2.

Desta forma, no exercício 1 o aluno relembrará o cálculo de uma distribuição binomial, enquanto que no exercício 2, os alunos compreenderão sobre a curva normal e também a consultar tabelas.

1 | ATIVIDADE 1

Primeiramente a atividade inciará com uma discussão envolvendo todos os alunos referente a variáveis discretas e contínuas.

Em seguida farão exercícios sobre o assunto, os quais servirão de motivação para a próxima atividade.

Concluídos os dois Exercícios, discutirão sobre a facilidade ou complexidade dos mesmos, expondo oralmente os motivos.

Espera-se que ao final da atividade os alunos realizem questionamentos, como por exemplo, se é possível resolver o Exercício 1 utilizando uma tabela, como no Exercício 2, consequentemente motivando-os para a próxima atividade.

Recursos: Calculadora (opcional) e a Tabela 2.

Duração: 2 horas-aula de 50 minutos cada.

Os Exercícios abaixo serão realizados pelos alunos.

O Exercício 1 foi extraído de [10], p. 196, conforme a questão:

1. Um estudo mostra que 60% dos casos de divórcio requeridos num certo município, a incompatibilidade é apontada como causa. Encontre as probabilidades de que entre 15 casos de divórcio requeridos naquele município

(a) no máximo cinco apontem a incompatibilidade como causa;

(b)de oito a onze apontem a incompatibilidade como causa;

(c)no mínimo onze apontem a incompatibilidade como causa.

O Exercício 2 foi extraído de [8], p. 208.

2. As contas de consumo mensais em uma cidade são normalmente distribuídas com uma média de R$100, 00 e um desvio padrão de R$12, 00. Uma conta de consumo é escolhida aleatoriamente.

(a) Encontre a probabilidade de a conta de consumo ser menor que R $70, 00;

(b) Encontre a probabilidade de a conta de consumo estar entre e R $90, 00 e R$120, 00;

(c) Encontre a probabilidade de a conta de consumo ser maior que R $140, 00.

2 | ATIVIDADE 2

O objetivo desta Atividade é mostrar, de maneira simples, que uma distribuição normal pode ser aproximada de uma distribuição binomial, e sob quais condições tal aproximação é possível.

Os recursos utilizados serão: Tábua de Galton (confeccionada pelo professor), bo- las de gude ou de rolamento em quantidade suficiente para cada aluno, calculadora (opcional) e a Tabela 2.

Para a concretização da atividade será apresentada a Tábua de Galton aos alunos, e também questionamento aos mesmos de onde é mais provável que a bola pare e onde é menos provável; solicitando que justifiquem as respostas apresentadas.

Em seguida, os alunos, um por vez, liberam as bolas, e verificam se suas hipóteses iniciais estavam ou não corretas.

Após realizada a atividade, o professor explica as probabilidades de cada bola nos primeiros pinos, deixando que os alunos façam as demais probabilidades.

Como forma de aguçar o raciocínio dos alunos serão realizadas as seguintes pergun- tas e atividades devem ser feitas:

- Podemos montar um gráfico da distribuição binomial? Como seria? Faça um gráfico que a represente.
- É possível formam um gráfico que indique a posição das bolas no experimento? Em caso afirmativo, este gráfico se aproxima do obtido anteriormente?
- Compare o gráfico com a curva normal, apresentada na Atividade anterior. O que podemos concluir?

Com tais exercícios espera-se que os alunos observem que existe uma relação entre as distribuições binomial e normal.

Após realizada esta fase, farão novamente o exercício 1 da atividade 1, desta vez através da distribuição normal.

Enfatizando que deve adicionar 0, 5 em cada extremidade, explicando-os que tal adição chama-se correção pela continuidade.

Assim, encerrada a atividade 2, os alunos serão questionados quanto aos resultados obtidos, observando se ocorreu diferenças ao realizado anteriormente e se é possível sempre fazer a aproximação.

Duração: 2 horas-aula de 50 minutos cada.

3 | ATIVIDADE 3

O objetivo desta Atividade é mostrar, de maneira informal, como obter uma amostra de uma proporção, além de mostrar que a média da proporção das amostras é igual a proporção da população.

Em primeiro lugar será feita uma indagação aos alunos sobre o que seria uma amostra como forma de observar os conhecimentos sobre o assunto. Em seguida apresentar a situação-problema através do enunciado:

Em um determinado município todas as famílias de um certo bairro têm animais de estimação, entre cães e gatos, cada família tem apenas um dos dois animais. Qual a proporção, neste bairro, de famílias que preferem ter cães como animais de estimação?

Exposto o enunciado o professor pergunta aos alunos o que seria mais fácil, contar todos os animais ou apenas uma parte dos animais?

Ouvidas as respostas, pede-se que os alunos marquem na folha, em um lugar aleatório, a quantidade de animais que eles quiserem, da maneira que preferirem, e contem apenas os animais que estão na área demarcada, anotando a quantidade de cães e gatos. O professor pode fazer a anotação caso alguma dupla assim prefira, como também pode fazer uso exemplos para demonstrar como a marcação pode ser feita.

Anotados os valores de todas as duplas, faz-se a média da proporção de cães, e compara-se com o valor exato, previamente conhecido pelo professor.

Em seguida, faz-se as seguintes perguntas aos alunos:

- O resultado está próximo da proporção real? Justifique.
- Qual amostra nos daria uma melhor estimativa, uma amostra pequena ou uma grande? Justifique.
- Nesta Atividade várias amostras foram obtidas, na prática trabalha-se com quan- tas amostras?

Recursos: Calculadora (opcional) e a Figura 45.

Duração: 1 hora-aula de 50 minutos.

4 I ATIVIDADE 4

O objetivo desta Atividade é obter a proporção de elementos com determinada característica de uma população a partir de uma única amostra, bem como obter um intervalo de confiança e verificar uma hipótese.

Será exposto a seguinte situação-problema:

Uma máquina fabrica 300 peças por hora, sendo que uma porcentagem destas peças apresenta algum defeito. O fabricante afirma que, para um intervalo de confiança de 95%, apenas 10% das peças são defeituosas. A afirmação do fabricante é verdadeira?

Para a realização desta atividade, será colocado em um saco plástico 300 peças, de duas cores distintas, as quais representam as peças fabricadas pela máquina, peças defeituosas na proporção 15% e peças que não possuem defeito na proporção 85%. As peças serão retiradas uma a uma com reposição.

A cada peça retirada anotamos a sua cor e a devolvemos para uma nova retirada. Ao final calculamos a proporção de peças de cada cor.

Após contagem e observação das peças serão feitas as seguintes perguntas:

- O que é um intervalo de confiança?
- O que é um intervalo de confiança de 95%?

Mostra-se como se efetua o cálculo de um intervalo de confiança de 95% para esta amostra com o uso da Tabela 2.

E na mesma continuidade mais questões serão apresentadas aos alunos:

- O que vem a ser uma hipótese?
- Qual é a hipótese a ser testada?
- A hipótese inicial pode ser aceita ou deve ser rejeitada? Justifique.

Comparando o resultado obtido com a proporção real, serão feitas as seguintes per- guntas:

- O resultado obtido é igual ou próximo ao real?
- A proporção real pertence ao intervalo de confiança?
- A aceitação (ou rejeição) da hipótese inicial foi uma decisão acertada?

Apresentado o teste para uma proporção e resolvendo-a, é verificado se a aceitação (ou rejeição) da hipótese inicial foi uma decisão acertada.

Recursos: Saco plástico ou sacola, 300 peças ou botões de duas cores distintas, cal- culadora (opcional) e Tabela 2.

Duração: 2 hora-aula de 50 minutos cada.

APLICAÇÃO DAS ATIVIDADES

Este capítulo destina-se à aplicação das Atividades anteriormente propostas, bem como as respostas dos alunos aos questionamentos, suas observações e fotos da realização das Atividades.

Os nomes dos alunos são fictícios e os cálculos foram efetuados com o auxílio de uma calculadora científica.

1 | APLICAÇÃO DA ATIVIDADE 1

Antes de se iniciar a atividade, houve o seguinte diálogo:

Professor: O que são variáveis aleatórias?

Marcelo: São coisas diferentes, mas ocorrem repetidamente.

Adriana: Uma coisa que varia de uma para outra.

Professor: De modo informal, variáveis aleatórias são os possíveis resultados de um evento que você espera que aconteçam, mas não sabe quando eles irão ocorrer.

Professor: E o que são variáveis aleatórias discretas? E variáveis aleatórias contínuas?

Adriana: Contínuas sempre se repetem.

Bianca: Discretas ocorrem de forma mais lenta.

Professor: Como assim, "mais lenta"?

Bianca: Demoram mais para acontecer do que as contínuas.

Renato: Contínuas seguem um padrão.

Professor: De modo mais formal, uma variável aleatória é uma função que assume valores no conjunto dos números reais. As variáveis aleatórias discretas assumem valores em um conjunto de pontos da reta e as variáveis aleatórias contínuas assumem valores em um intervalo no conjunto dos números reais.

Os alunos foram divididos em grupos escolhidos por eles e foi apresentado o Exercício 1:

Professor: Lembram dos exercícios do ano passado, de sucesso e fracasso?

Alunos: Sim.

Professor: Neste caso, a incompatibilidade é sucesso ou fracasso?

Vários alunos: Fracasso.

Professor: Mas, o que é esperado que ocorra?

Marcelo: A incompatibilidade.

Professor: Então, a incompatibilidade é sucesso ou fracasso?

Marcelo: Sucesso.

Os demais concordaram com a resposta de Marcelo, e seguimos a análise do exercício.

Professor: Qual o significado de no máximo 5 no item (*a*)?

Bianca: Mais do que 5.

Carla: 5.

Professor: Exatamente 5?

Carla: Sim.

Adriana: Menores do que 5, o 5 também faz parte.

Professor: Isso mesmo. E o zero?

Marcelo: O zero também.

Escrevi no quadro a conta a ser feita, com a participação dos alunos, e estes auxiliaram na resolução do item (*a*), cada grupo fez uma parte, compartilhando a mesma calculadora.

No item (*b*) Michele perguntou se poderia ir do oito ao onze direto.

Professor: Precisamos fazer a conta com o nove e o dez.

Neste item os alunos chegaram ao resultado com orientação e Márcio indagou se contínuas estão no intervalo.

Professor: O que seria no mínimo 11 no item (*c*)?

Carla: Não pode ser mais do que 11.

Adriana: Devem ser de 11 "pra" cima. **Professor:** E o 11?

Adriana: O 11 também.

Alguns alunos disseram que a conta era difícil, outros que era trabalhosa, complicada, confusa.

Mesmo assim, tentaram encontrar o resultado novamente, pois a calculadora que usaram inicialmente faz a combinação direto e eles queriam usar a calculadora deles, com o fatorial, e conseguiram obter os resultados corretos.

Seguem os cálculos efetuados no Exercício 1:

(a) $\binom{15}{0}(0,6)^0(0,4)^{15} + \binom{15}{1}(0,6)^1(0,4)^{14} + \binom{15}{2}(0,6)^2(0,4)^{13} + \binom{15}{3}(0,6)^3(0,4)^{12} +$
$+\binom{15}{4}(0,6)^4(0,4)^{11} + \binom{15}{5}(0,6)^5(0,4)^{10} = 0,034.$

(b) $\binom{15}{8}(0,6)^8(0,4)^7 + \binom{15}{9}(0,6)^9(0,4)^6 + \binom{15}{10}(0,6)^{10}(0,4)^5 + \binom{15}{11}(0,6)^{11}(0,4)^4 = 0,696.$

(c) $\binom{15}{11}(0,6)^{11}(0,4)^4 + \binom{15}{12}(0,6)^{12}(0,4)^3 + \binom{15}{13}(0,6)^{13}(0,4)^2 + \binom{15}{14}(0,6)^{14}(0,4)^1 +$
$+\binom{15}{15}(0,6)^{15}(0,4)^0 = 0,217.$

O Exercício 2 é apresentado e diversas perguntas são feitas aos alunos:

Professor: O que significa "normalmente distribuídas"?

Michele: São coisas distribuídas igualmente.

Professor: Não, não é isso.

Marcelo: São coisas distribuídas de forma diferente.

Professor: Também não. Quando se diz que algo é normalmente distribuído, é porque ele segue uma curva normal.

Faço um esboço de uma curva normal no quadro.

Professor: No exercício anterior, a variável aleatória era discreta.

Carla: Por que?

Professor: Porque os valores são números inteiros. Temos 8 casais e meio?

Marcelo: Não.

Professor: Agora, uma conta de consumo pode ser de R$ 117, 35?

Marcelo: Pode.

Professor: Então, nesse exercício, a variável aleatória é contínua?

Adriana: É.

Professor: O que é uma média 100?

Bianca: Mais ou menos 100.

Professor: Se somarmos todas as contas e dividirmos esse resultado pelo total de contas encontramos cem, que é a média do exercício. E em que lugar do gráfico (apontando para a reta abaixo da curva) vai estar a média cem?

Carla: Bem no meio da curva.

Professor: E o que é um desvio padrão?

Bianca: É um desvio que ocorre regulamente.

Professor: O desvio padrão indica o quanto um conjunto de valores se afasta da média. Por exemplo, um desvio padrão igual a doze, indica que 68% dos valores das contas de consumo pertencem ao intervalo entre 88 e 112.

Com o item (*a*) surgiram muitas dúvidas por parte dos alunos.

Michele: Quando o número fica para a esquerda da média é negativo? (referindo-se à Tabela da distribuição normal).

Professor: Na tabela sim.

Adriana: Existe uma tabela que tenha todos os valores positivos?

Professor: Existe sim.

Adriana: Então por que você usou essa?

Professor: Eu tinha que fazer uma escolha, e optei por essa.

Michele: Desvio padrão interfere no número ser negativo ou positivo?

Professor: Não interfere, quando fizermos os cálculos você verá.

Michele: Posso colocar qualquer valor abaixo da média? (perguntando se podia escolher qualquer valor que fosse menor do que 100).

Professor: Pode sim, pode ser menor, maior ou igual a 100.

Carla: Por que a média é 100?

Professor: É o valor fornecido pelo problema, por exemplo, se você somar todas as contas de consumo e dividir pelo número total de contas, você chegaria a cem, neste caso.

Foi demonstrado como fazer o cálculo das probabilidade utilizando a Tabela de distribuição normal (esse assunto era novidade para eles).

Carla: O que se faz com o número negativo? Passa para positivo?

Professor: Tem uma observação ao final da tabela.

Na folha fornecida com a Tabela de distribuição normal tinha a seguinte observação:

$$\Phi(-x) = 1 - \Phi(x).$$

Ester: Então, com o sinal negativo fazemos 1 menos o número só que agora positivo e procuramos na tabela.

Os alunos tiveram uma certa dificuldade em fazer a subtração $1 - 0,9938$. Quando mal havia terminado de explicar eles conseguiram terminar sem auxílio.

Iniciado o item (*b*) Carla perguntou novamente o porquê do cem sempre aparecer.

Ester: A média é 100, sempre iremos usar 100.

Professor: Somente neste exercício, pois a média é 100.

Ester: Isso, só nesse exercício.

No item (*b*) obtivemos $\Phi(1,67) - \Phi(-0,83)$ e novos questionamentos surgiram.

Ester: O valor está na linha do 1, 6, não é professor?

Professor: Sim, mas é na 8a coluna de valores.

Adriana: Por que na 8a coluna é que se encontra o valor e não em outra?

Professor: Olhem a 1a linha da tabela, elas representam os centésimos, temos que fazer $1,6 + 0,07 = 1,67$.

Ester: Então, 0, 83 está na 4a coluna?

Professor: Isso mesmo.

Este item foi concluído com mais facilidade que o anterior.

Professor: No item (*c*) temos que fazer $1 - \Phi(3,33)$.

Marcelo: Por que, se ele é positivo?

Professor: É porque a tabela nos fornece valores à esquerda de $\Phi(3,33)$, isso quer dizer que nos fornece a área dos valores menores do que R$ 140, 00, e queremos a área dos valores maiores do que R$ 140, 00. Como a área toda é igual a 1, fazemos $1 - \Phi(3,33)$.

Marcelo: Entendi.

Os alunos tiveram mais facilidade com este exercício, em comparação ao anterior. Seguem os cálculos efetuados no Exercício 2:

(a)

$$\begin{aligned} P\left(x < \frac{70-100}{12}\right) &= P(x < -2,5) \\ &= \Phi(-2,5) \\ &= 1 - \Phi(2,5) \\ &= 1 - 0,9938 \\ &= 0,0062. \end{aligned}$$

(b)

$$\begin{aligned} P\left(\frac{90-100}{12} \leqslant x \leqslant \frac{120-100}{12}\right) &= P\left(-0,83 \leqslant x \leqslant 1,67\right) \\ &= \Phi(1,67) - \Phi(-0,83) \\ &= \Phi(1,67) - [1 - \Phi(0,83)] \\ &= 0,9525 - 1 + 0,7967 \\ &= 0,7492. \end{aligned}$$

(c)

$$\begin{aligned} P\left(\frac{140-100}{12} < x\right) &= P\left(3,33 < x\right) \\ &= 1 - \Phi(3,33) \\ &= 1 - 0,9996 \\ &= 0,0004. \end{aligned}$$

Professor: Será que podemos fazer o cálculo do exercício 1 como fizemos no exercício 2?

Carla: Com certeza.

Marcelo: Eu acho que sim, mas esse ainda é complicado.

Professor: É que foi a primeira vez que vocês fizeram, as próximas serão menos complicadas.

2 I APLICAÇÃO DA ATIVIDADE 2

Esta Atividade começou antes de ser levada para a sala de aula, pela construção da Tábua de Galton.

Em um primeiro momento foi feita uma construção geométrica com diversos triângulo equiláteros, cujos vértices seriam os pontos no qual seriam colocados os pregos, esta figura foi presa à tábua de madeira com fita adesiva, Figura 16.

Figura 16: Construção da Tábua de Galton 1

Fonte: Arquivo próprio do autor.

Foram colocados os pregos na tábua, Figura 17.
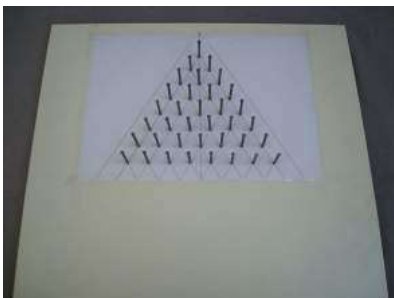
Figura 17: Construção da Tábua de Galton 2

Fonte: Arquivo próprio do autor.

As linhas laterais eram destinadas a colocar uma ripa de madeira com o intuito de que as bolas de gude não escapassem, mas não funcionou muito bem, sendo necessário a colocação de mais pregos, Figura 18.
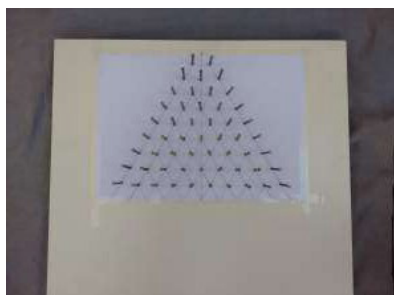
Figura 18: Construção da Tábua de Galton 3

Fonte: Arquivo próprio do autor.

Por fim, foram coladas as canaletas onde ficariam as bolas de gude, Figura 19.
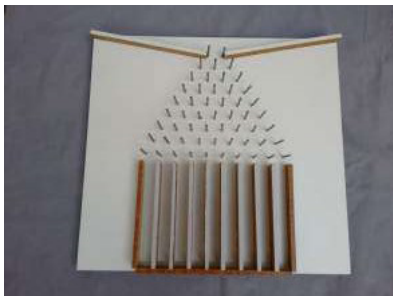
Figura 19: Construção da Tábua de Galton 4

Fonte: Arquivo próprio do autor.

Após a confecção da tábua de Galton foi levada para a sala de aula, o que criou grande curiosidade entre os alunos, sobre o que seria e qual a sua utilidade. Houve a explicação sobre o que seria, mas sua utilidade não, pelo menos em um primeiro momento.

Professor: Esta é a Tábua de Galton, quando colocarem uma bola de gude na parte de cima ela cairá e baterá nos pregos, chegando em uma das canaletas. Onde vocês acham que a bolinha cairá?

Adriana e Bianca: no meio.

Carla: Na ponta.

Marcelo: Na 3a canaleta.

Marcelo: Podemos numerar de 1 até 9?

Professor: Assim ficaria mais fácil?

Marcelo: Sim.

Professor: Todos concordam, podemos numerar as casas de 1 até 9?

Todos: Sim.

Lançada a primeira bola, ela cai no canto, como Carla havia dito anteriormente.

Figura 20: Primeira bola lançada

Fonte: Arquivo próprio do autor.

Carla: Não falei que ia cair no canto.

Mas as 3 bolas seguintes caíram todas em um mesmo lugar e Carla já não parecia tão certa de sua escolha.
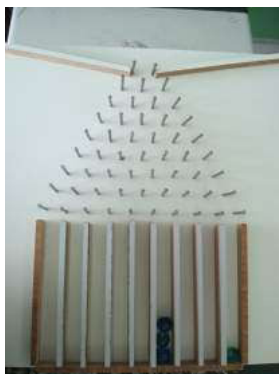
Figura 21: Quatro bolas lançadas

Fonte: Arquivo próprio do autor.

Guilherme: Mas só vai cair aí?

Professor: Calma, vamos colocar mais bolas.

Mais bolas foram colocadas, até que atingimos o limite em uma das canaletas, num total de 32 bolas. Figura 23.
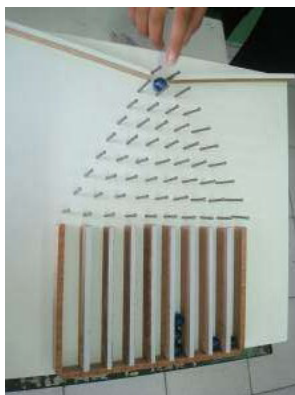
Figura 22: Lançando uma bola

Fonte: Arquivo próprio do autor.
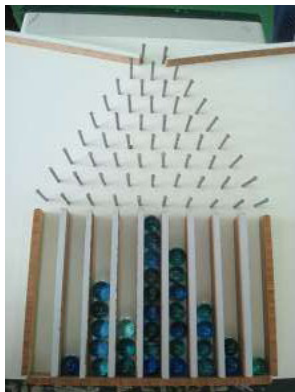
Figura 23: Final do experimento, 32 bolas lançadas

Fonte: Arquivo próprio do autor.

Ester: Caiu mais no meio, por que?

Comecei explicando as probabilidades na própria Tábua, Figura 24.

Professor: Quando a bola bate no primeiro prego, qual a probabilidade de que ela vá para a esquerda?

Marcelo: É de 50%.

Professor: E para a direita?

Marcelo: É de 50%.

Professor: Se a bola for para a direita ela bate no prego de baixo. Qual a probabilidade dela ir para a esquerda?

Ester: É de 50%.

Professor: Não é mais 50%, agora a probabilidade é de 25%.

Carla: Por que 25%?

Professor: A bola não tinha 50% de chance de ir para a esquerda?

Ester: Tinha.

Professor: Agora, só temos 50% de chance para a bola escolher um dos lados, 25% de ir para a esquerda e 25% de ir para a direita. E se a bola for para a direta?

Ester: A mesma coisa, 25% para cada lado.

Professor: Mas a bola não pode ir para o meio vindo de qualquer um dos lados?

Marcelo: Pode.

Professor: Então, ela tem 25% de chance de vir da esquerda e 25% de chance de vir da direita?

Marcelo: Tem.

Professor: Isso significa que ela tem 50% de chance de cair no meio.

Marcelo: Sim.

Ester: Isso quer dizer que ela vai cair no meio?

Professor: Não necessariamente, porém a maior probabilidade é de que ela caia no meio.

Ester: Por isso é que no meio caiu mais.

Professor: Isso mesmo.

Comecei fazendo as probabilidade no quadro, Figura 25. A sala foi dividida em dois grupos, um grupo fez um gráfico da posição das bolas na Tábua e o outro as probabilidades iniciadas no quadro.

No início, os alunos fizeram livremente os gráficos.

Figura 24: Explicando as probabilidades na Tábua

Fonte: Arquivo próprio do autor.
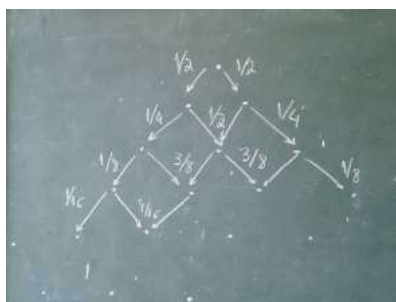
Figura 25: Probabilidades no quadro

Fonte: Arquivo próprio do autor.
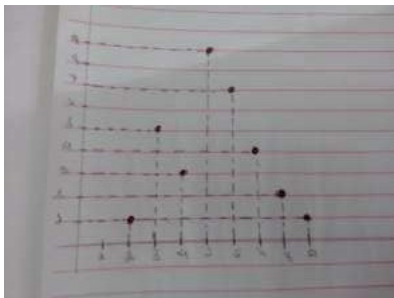
Figura 26: Gráfico da Carla

Fonte: Arquivo próprio do autor.

Depois, foi explicado que seria melhor fazer um histograma, pois ele melhor representa a distribuição.

Ester: Como se faz um histograma?

Professor: É como o gráfico que a Marta fez, mas as colunas são todas coladas umas nas outras.
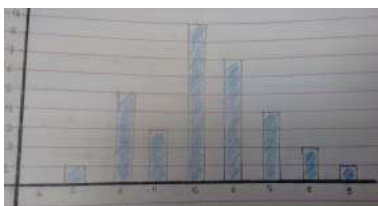
Figura 27: Primeiro gráfico da Marta

Fonte: Arquivo próprio do autor.

Marta: Assim está bom?

Professor: Está ótimo.
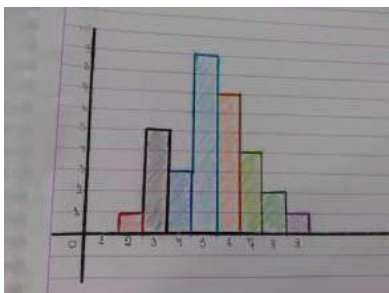
Figura 28: Segundo gráfico da Marta

Fonte: Arquivo próprio do autor.

Os alunos responsáveis pelas probabilidades tiveram dificuldades em concluir a tarefa. Adriana foi uma das poucas pessoas a concluir a tarefa.
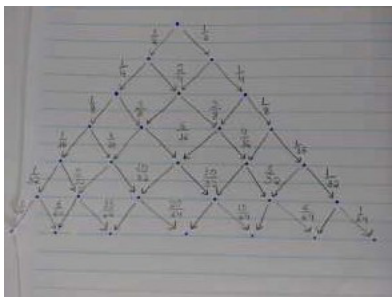
Figura 29: Probabilidades da Adriana

Fonte: Arquivo próprio do autor.

Como houve muita dificuldade resolvi explicar no quadro.

Professor: Lembram do triângulo de Pascal, estudado no ano anterior?

Pedro: Lembro.

Professor: O numerador segue o triângulo de Pascal. E o denominador, saberia responder?

Adriana: A probabilidade em cada linha.

Carla: Por que 16? (referindo-se ao denominador)

Professor: A metade de 1⁄8 é 1⁄16.

Carla: Não entendi.

Professor: Lembra como dividimos uma fração por outra?

Carla: Não.

Através de um desenho no quadro, Carla entendeu.

Figura 30: Explicando as Probabilidades

Fonte: Arquivo próprio do autor.

Ester: Então 20⁄64 o é a probabilidade de cair no meio?

Professor: Isso mesmo.

Ester: Mas o que significa o 64?

Professor: Significa que se colocarmos 64 bolas, espera-se que 20 caiam no meio.

Marcelo: E se não cair?

Professor: Exatamente 20?

Marcelo: É.

Professor: Não tem problema, 20 é o resultado esperado. Podem cair mais ou menos do que 20 bolas.

Ester: Isso significa que sempre vai cair mais no meio?

Professor: Espera-se que sim.

Foi solicitado que os alunos fizessem um gráfico das probabilidades que estavam no quadro.
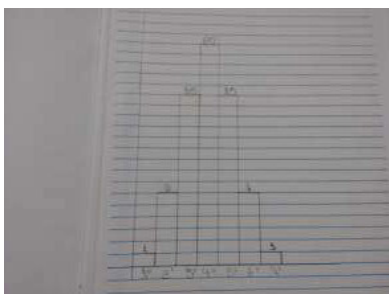
Figura 31: Gráfico do Marcelo

Fonte: Arquivo próprio do autor.

Após comparar os gráficos da distribuição binomial e o obtido na Tábua com a curva normal, houve concordância por parte dos alunos que ambos se aproximavam da curva normal.

Resolveram novamente o Exercício 1, desta vez pela aproximação da binomial à normal e a maior dificuldade dos alunos foi na compreensão da correção pela continuidade.

Professor: Como o cálculo é uma aproximação da binomial precisamos encontrar uma média.

Adriana: Como encontramos a média?

Professor: Qual a probabilidade de sucesso?

Adriana: É de 60%.

Professor: E em quantos casais estamos interessados?

Marcelo: Em 15.

Professor: A média é igual a 15 multiplicado por 0, 6, sendo igual a 9. Calculamos o desvio padrão, encontrando 1, 8974.

Professor: Temos mais um problema, não podemos utilizar o 5 no item (*a*).

Marcelo: Por que?

Professor: Se observarmos o gráfico, o 5 está localizado bem no meio da coluna, se usarmos somente o 5, fica faltando metade da coluna, e a coluna toda está no intervalo entre 4, 5 e 5, 5, por isso usamos o 5, 5.

Adriana: E porque não podemos usar o 4, 5?

Professor: Porque o 5 ficaria de fora do cálculo, e queremos calcular no máximo 5. E no item (*b*)?

Ester: Temos que fazer como antes. (Referindo-se ao item (*a*)).

Professor: Mas utilizamos 7, 5 ou 8, 5? 10, 5 ou 11, 5?

Ester: O 8, 5 e o 11, 5.

Professor: Não, um deles não está correto.

Adriana: É 7, 5?

Professor: Isso mesmo. Por quê?

Adriana: O 8 ficaria de fora.

Através de um esboço, a explicação do porquê utiliza-se o intervalo entre 7, 5 e 11, 5; foi mais clara.

Professor: E no item (*c*)?

Ester: Usamos o 10, 5.

Professor: Por quê?

Ester: Lembrei do desenho da aula anterior. Fica tudo para a direita, se fosse 11, 5 o 11 ficaria de fora.

Os alunos não apresentaram grandes dificuldades nos cálculos, eles estavam mais familiarizados com a Tabela, logo encontraram os valores com certa facilidade.

Seguem os cálculos efetuados:

(a)

$$\begin{aligned} P\left(x < \frac{5,5-9}{1,8974}\right) &= P\left(x < -1,85\right) \\ &= \Phi(-1,85) \\ &= 1 - \Phi(1,85) \\ &= 1 - 0,9678 \\ &= 0,0322. \end{aligned}$$

(b)

$$\begin{aligned} P\left(\frac{7,5-9}{1,8974} \leqslant x \leqslant \frac{11,5-9}{1,8974}\right) &= P\left(-0,79 \leqslant x \leqslant 1,32\right) \\ &= \Phi(1,32) - \Phi(-0,79) \\ &= \Phi(1,32) - [1 - \Phi(0,79)] \end{aligned}$$

$$= 0,9066 - 1 + 0,7852$$
$$= 0,6918.$$

(c)

$$\begin{aligned} P\left(\frac{10,5-9}{1,8974} < x\right) &= P(0,79 < x) \\ &= 1 - \Phi(0,79) \\ &= 1 - 0,7852 \\ &= 0,2148. \end{aligned}$$

Professor: Os resultado encontrados, se comparados com os anteriores, podem ser considerados boas aproximações?

Vários alunos: Sim.

Professor: Sempre é possível fazer a aproximação?

Nícolas: Sim.

Muitos alunos concordaram com a resposta do Nícolas, então foi mostrado sob que condições é possível fazer a aproximação.

Ester: Por que não podemos fazer se forem menores do que 5?

Professor: Os resultados podem não ser confiáveis.

Adriana: Os resultados anteriores são exatos?

Professor: Sim, são exatos. Então, qual a vantagem em fazer aproximações?

Ester: Fazer menos contas.

Professor: Seria o de reduzir o tempo com cálculos excessivos, quando uma boa aproximação já é suficiente.

3 | APLICAÇÃO DA ATIVIDADE 3

Esta atividade teve início com a seguinte pergunta:

Professor: O que seria uma amostra?

Fabíola: Uma parte de alguma coisa, por exemplo, uma amostra de xampu.

Marcelo: Ou uma amostra de perfume.

Professor: Temos uma amostra quando queremos fazer uma pesquisa sobre algo e obtemos informações de uma parte da população, como por exemplo, uma pesquisa eleitoral não entrevista toda a população de uma cidade ou de um país, pois tomaria muito tempo e seria muita caro.

Os alunos foram divididos em duplas, alguns pediram para ser um trio, logo havia duplas e trios. Foram distribuídas as folhas destinadas a esta atividade e explicada a situação-problema.

Solicitando que selecionassem uma amostra da maneira que quisessem. No início

ficaram desorientados, assim foi necessário apresentar um exemplo de como poderiam obter uma amostra.

Figura 32: Exemplo do Professor

Fonte: Arquivo próprio do autor.

Os alunos se empenharam nessa atividade.

As proporções encontradas pelos grupos foram:

0, 76 ; 0, 72 ; 0, 69 ; 0, 62 ; 0, 72 ; 0, 71 ; 0, 87 ; 0, 73 e 0, 58.

Figura 33: Participação dos alunos 1

Fonte: Arquivo próprio do autor.

Figura 34: Participação dos alunos 2

Fonte: Arquivo próprio do autor.

A soma feita por Bianca foi igual a 6, 4.

Este resultado foi dividido pela quantidade de grupos, onde obteve-se 0, 71.

A quantidade real era de 248 cães e 136 gatos, num total de 384 animais. Portanto, 64, 58% dos animais eram cães e 35, 42% eram gatos.

Professor: Como podemos considerar a nossa aproximação?

Marcelo: Boa.

Professor: Por quê?

Marcelo: Chegamos a um resultado muito próximo do real. A marcação de Adriana me chamou a atenção.

Professor: Por que você só fez uma linha?

Adriana: É porque eu gosto mais de cachorros.

Professor: Mas você percebeu que isso influenciou o resultado?

Adriana: Comparando com os outros resultados, sim.

Professor: Em uma pesquisa temos que ser imparciais, senão toda a pesquisa pode ser comprometida.

Angélica: Por que alguns grupos contaram a mesma quantidade de animais mas a quantidade de cachorros foi diferente?

Professor: Isso depende da região na qual esses grupos obtiveram seus resultados.

Professor: Coletamos ao todo 9 amostras, em uma pesquisa real, quantas amostras são coletadas?

Nícolas: 10.

Bianca: 20.

Professor: Elas não podem ser muitas, pois a pesquisa tomaria muito tempo e seria muito cara. Geralmente trabalhamos com uma única amostra.

A seguir seguem as marcações de alguns grupos:

Figura 35: Grupo do Marcelo

Fonte: Arquivo próprio do autor.

4 I APLICAÇÃO DA ATIVIDADE 4

Antes de iniciar a Atividade, explicou-se que haveria um sorteio e que iríamos obter uma amostragem aleatória simples.

Professor: O que é uma amostragem?

Camila: É uma amostra de algo.
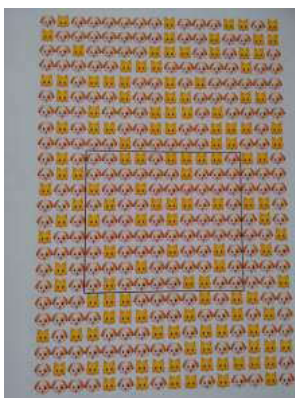
Figura 36: Grupo da Ester

Fonte: Arquivo próprio do autor.
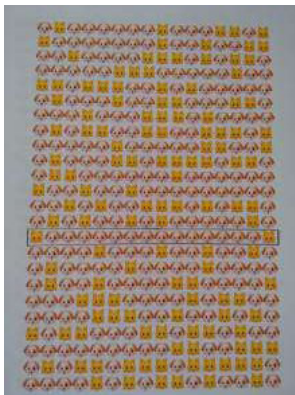
Figura 37: Grupo da Adriana

Fonte: Arquivo próprio do autor.

Professor: O que é uma amostragem aleatória simples?

Pedro: É uma amostra que você não sabe o que vai obter.

Professor: Em uma amostragem aleatória simples todos o elementos tem a mesma chance de ser sorteados.

Adriana: E como se faz isso?

Professor: Sempre que uma peça for sorteada ela é devolvida para que a probabilidade seja sempre a mesma. Chamamos de retirada com reposição.

Marcelo: E como a gente sabe o que foi sorteado?

Professor: Anotamos os resultados.

Explicada a situação-problema, os alunos foram orientados a considerar que os botões escuros representariam as peças não defeituosas enquanto os botões claros representariam as peças defeituosas.

Figura 38: Botões representando peças

Fonte: Arquivo próprio do autor.

Marcelo, Ester e Carla fizeram o sorteio com todos os alunos enquanto, anotava os resultados no quadro. Foi feito o sorteio de 40 peças, apesar de haver 32 alunos na sala.

Figura 39: Sorteio de uma peça

Fonte: Arquivo próprio do autor.

No primeiro sorteio obtivemos uma peça escura.

O resultado foi anotado no quadro.

A primeira peça clara apareceu no sorteio da 7o peça.

Enquanto as peças eram sorteadas, alguns alunos fizeram comentários:

Karina: Tem mais botões escuros? (Ela fez este comentário três vezes).

Adriana: Eu tirei um botão escuro, eu não sou defeituosa.

Figura 40: Sorteio de uma peça escura

Fonte: Arquivo próprio do autor.

Figura 41: Anotando o resultado

Fonte: Arquivo próprio do autor.

Figura 42: Sorteio de uma peça clara

Fonte: Arquivo próprio do autor.

Após as 40 retiradas foi verificado que 35 botões eram escuros e 5 eram claros.

Professor: O que é intervalo de confiança?

Marcelo: Quando se tem noção que a peça está em boas condições.

Karina: Quando você acredita que as peças estão boas.

Figura 43: Resultado do sorteio

Fonte: Arquivo próprio do autor.

Professor: E o que é um intervalo de confiança de 95%?

Marcelo: Que 95% das peças não apresenta vício.

Professor: Um intervalo de confiança de 95% é um intervalo em que você confia que a sua afirmação está correta em 95% das amostras.

Apresentado o cálculo do intervalo de confiança, explicando que

$$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

seria o erro máximo cometido, n o número de retiradas e $\hat{p}$ a proporção de peças defeituosas na amostra, $z_{\alpha/2}$ explicando com mais detalhes.

Professor: Mas, antes de iniciar temos algo muito importante para verificar.

Ester: O quê?

Professor: Se podemos fazer a aproximação pela normal.

Marcelo: Podemos.

Professor: Como você sabe?

Marcelo: Fiz as contas.

Professor: Quais contas?

Marcelo: Fiz $0,125 \cdot 40 = 5$ e $0,875 \cdot 40 = 35$.

Professor: O intervalo de confiança é $1 - \alpha$ é igual a 0, 95, então α é igual a...

Ester: α é igual a 0, 05.

Professor: E $\alpha/2$?

Ester: É igual a 0, 025.

Professor: Vamos fazer uma figura que represente o que acabamos de falar.

A figura em questão é a Figura 44.

Figura 44: Intervalo de confiança

Fonte: Arquivo próprio do autor.

Professor: A nossa tabela serve para os valores à direita do centro. Então temos que fazer 1 – 0, 025.

Adriana: Que é igual a 0, 975.

Professor: Isso. Agora procurem 0, 975 na tabela.

Renato foi o primeiro a encontrar o valor correto de 1, 96, logo todos encontraram o valor correto, demonstrando mais habilidade em manusear a tabela.

Obtemos o intervalo de confiança a seguir:

$$\left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] = [0,0225; 0,2275] .$$

Os alunos fizeram os cálculos com o auxílio de uma calculadora e orientação do professor.

Professor: O que vem a ser uma hipótese?

Marcelo: É uma coisa que você não sabe se é verdade e quer testar para saber se é mesmo verdade.

Professor: Qual a hipótese a ser testada?

Adriana: Que 10% das peças serem defeituosas é verdade.

Professor: A hipótese inicial deve ser aceita ou rejeitada?

Adriana: Aceita.

Professor: Por quê?

Adriana: Ela está dentro do intervalo de confiança.

Professor: A quantidade de botões é de 45 claros e 255 escuros, 15% dos botões são claros. Este valor está próximo do encontrado?

Ester: Sim.

Professor: E está dentro do intervalo de confiança?

Adriana: Está.

Professor: Pelo que vimos, a decisão de aceitar a hipótese foi uma decisão acertada?

Marcelo: Sim, foi, porque está dentro do intervalo de confiança.

Professor: Precisamos fazer um teste para descobrir.

Ester: Qual teste.

Foi apresentado o teste a seguir e seu resultado.

$$z = \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1-\hat{p})/n}} = \frac{0,125 - 0,1}{\sqrt{0,125 \cdot 0,875/40}} = 0,478.$$

Adriana: O que é p_0?

Professor: É o valor que o fabricante informou. Como o resultado está entre –1, 96 e 1, 96 devemos aceitar a hipótese como verdadeira.

Marcelo: Então acertamos.

CONSIDERAÇÕES FINAIS

Ao pensar um determinado tema para escrever a conclusão do curso, queria um assunto não exclusivamente para usar em sala de aula, nas aulas de matemática, mas sim um assunto que extrapolasse os muros da escola, e que contribuísse de forma significativa na vida em sociedade dos jovens educandos, como também a utilidade em outras áreas do conhecimento. Desta forma, a estatística foi escolhida, justamente por cumprir as expectativas desejadas.

Logo surgiram as situação-problema entre elas: como seria se algum produto de certa empresa onde uma pequena porcentagem dos itens produzidos por uma determinada máquina tivesse algum defeito e esta empresa quisesse saber qual a porcentagem destes itens, sem perder muito tempo verificando-os individualmente.

O principal objetivo seria verificar qual o método estatístico apropriado e de maior confiabilidade para tal situação e principalmente mostrar aos alunos do Ensino Médio de que forma é realizada uma estimativa a partir de uma amostra e porque ela funciona; como também verificar se o tamanho da amostra tem influência no resultado obtido.

No entanto, em meados de 2015, após a leitura sobre a história da estatística, o foco mudou completamente, então surgiu o interesse pela inferência estatística, o qual considerei um assunto extremamente mais atrativo.

E consequentemente surgiu a ideia de construir a Tábua de Galton, sendo necessário primeiramente uma atividade introdutória; a atividade foi considerada "maçante" pelos alunos; porque envolveu uma grande quantidade de cálculos; diferentemente das posteriores atividades desenvolvidas.

A partir da segunda atividade, os alunos utilizaram a Tábua, despertando mais atenção em todos, pois queriam ver as bolas batendo nos pregos e caindo nas canaletas.

O resultado foi satisfatório visto que aproximou-se muito da curva normal, principalmente o lado direito.

Na terceira atividade, o resultado obtido ficou muito próximo do real, porém poderia ser bem melhor, se as escolhas escolhas dos locais ocorressem de forma mais aleatória, a pouca quantidade de exemplos influenciou a escolha dos alunos, e os resultados obtidos na quarta atividade foram muito próximo do real.

Ao término das diversas atividades, exercícios, conversas, explicações, cálculos; conclui- se que os resultados obtidos foram excelentes; principalmente porque todos os alunos desenvolveram várias habilidades, entre elas compreenderam a importância de uma pesquisa estatística e percebiam o impacto de uma única amostra na pesquisa.

Ao longo do trabalho algumas dificuldades surgiram, em especial a questão do tempo para a aplicação das atividades, pois foram duas aulas semanais e o prazo para o termino do desenvolvimento das atividades.

Outro aspecto relevante foi que a estatística é tema do final do 3o ano, sendo que

as atividades foram trabalhadas no início do ano letivo, entre os meses de março e abril, ou seja, os alunos ainda não possuíam completo domínio do assunto.

TABELAS

X	0, 00	0, 01	0, 02	0, 03	0, 04	0, 05	0, 06	0, 07	0, 08	0, 09
0, 0	0, 5000	0, 5040	0, 5080	0, 5120	0, 5160	0, 5199	0, 5239	0, 5279	0, 5319	0, 5359
0, 1	0, 5398	0, 5438	0, 5478	0, 5517	0, 5557	0, 5596	0, 5636	0, 5675	0, 5714	0, 5753
0, 2	0, 5793	0, 5832	0, 5871	0, 5910	0, 5948	0, 5987	0, 6026	0, 6064	0, 6103	0, 6141
0, 3	0, 6179	0, 6217	0, 6255	0, 6293	0, 6331	0, 6368	0, 6406	0, 6443	0, 6480	0, 6517
0, 4	0, 6554	0, 6591	0, 6628	0, 6664	0, 6700	0, 6736	0, 6772	0, 6808	0, 6844	0, 6879
0, 5	0, 6915	0, 6950	0, 6985	0, 7019	0, 7054	0, 7088	0, 7123	0, 7157	0, 7190	0, 7224
0, 6	0, 7257	0, 7291	0, 7324	0, 7357	0, 7389	0, 7422	0, 7454	0, 7486	0, 7517	0, 7549
0, 7	0, 7580	0, 7611	0, 7642	0, 7673	0, 7704	0, 7734	0, 7764	0, 7794	0, 7823	0, 7852
0, 8	0, 7881	0, 7910	0, 7939	0, 7967	0, 7995	0, 8023	0, 8051	0, 8078	0, 8106	0, 8133
0, 9	0, 8159	0, 8186	0, 8212	0, 8238	0, 8264	0, 8289	0, 8315	0, 8340	0, 8365	0, 8389
1, 0	0, 8413	0, 8438	0, 8461	0, 8485	0, 8508	0, 8531	0, 8554	0, 8577	0, 8599	0, 8621
1, 1	0, 8643	0, 8665	0, 8686	0, 8708	0, 8729	0, 8749	0, 8770	0, 8790	0, 8810	0, 8830
1, 2	0, 8849	0, 8869	0, 8888	0, 8907	0, 8925	0, 8944	0, 8962	0, 8980	0, 8997	0, 9015
1, 3	0, 9032	0, 9049	0, 9066	0, 9082	0, 9099	0, 9115	0, 9131	0, 9147	0, 9162	0, 9177
1, 4	0, 9192	0, 9207	0, 9222	0, 9236	0, 9251	0, 9265	0, 9279	0, 9292	0, 9306	0, 9319
1, 5	0, 9332	0, 9345	0, 9357	0, 9370	0, 9382	0, 9394	0, 9406	0, 9418	0, 9429	0, 9441
1, 6	0, 9452	0, 9463	0, 9474	0, 9484	0, 9495	0, 9505	0, 9515	0, 9525	0, 9535	0, 9545
1, 7	0, 9554	0, 9564	0, 9573	0, 9582	0, 9591	0, 9599	0, 9608	0, 9616	0, 9625	0, 9633
1, 8	0, 9641	0, 9649	0, 9656	0, 9664	0, 9671	0, 9678	0, 9686	0, 9693	0, 9699	0, 9706
1, 9	0, 9713	0, 9719	0, 9726	0, 9732	0, 9738	0, 9744	0, 9750	0, 9756	0, 9761	0, 9767
2, 0	0, 9772	0, 9778	0, 9783	0, 9788	0, 9793	0, 9798	0, 9803	0, 9808	0, 9812	0, 9817
2, 1	0, 9821	0, 9826	0, 9830	0, 9834	0, 9838	0, 9842	0, 9846	0, 9850	0, 9854	0, 9857
2, 2	0, 9861	0, 9864	0, 9868	0, 9871	0, 9875	0, 9878	0, 9881	0, 9884	0, 9887	0, 9890
2, 3	0, 9893	0, 9896	0, 9898	0, 9901	0, 9904	0, 9906	0, 9909	0, 9911	0, 9913	0, 9916
2, 4	0, 9918	0, 9920	0, 9922	0, 9925	0, 9927	0, 9929	0, 9931	0, 9932	0, 9934	0, 9936
2, 5	0, 9938	0, 9940	0.9941	0, 9943	0, 9945	0, 9946	0, 9948	0, 9949	0, 9951	0, 9952
2, 6	0, 9953	0, 9955	0, 9956	0, 9957	0, 9959	0, 9960	0, 9961	0, 9962	0, 9963	0, 9964
2, 7	0, 9965	0, 9966	0, 9967	0, 9968	0, 9969	0, 9970	0, 9971	0, 9972	0, 9973	0, 9974
2, 8	0, 9974	0, 9975	0, 9976	0, 9977	0, 9977	0, 9978	0, 9979	0, 9979	0, 9980	0, 9981
2, 9	0, 9981	0, 9982	0, 9982	0, 9983	0, 9984	0, 9984	0, 9985	0, 9985	0, 9986	0, 9986
3, 0	0, 9987	0, 9987	0, 9987	0, 9988	0, 9988	0, 9989	0, 9989	0, 9989	0, 9990	0, 9990
3, 1	0, 9990	0, 9991	0, 9991	0, 9991	0, 9992	0, 9992	0, 9992	0, 9992	0, 9993	0, 9993
3, 2	0, 9993	0, 9993	0, 9994	0, 9994	0, 9994	0, 9994	0, 9994	0, 9995	0, 9995	0, 9995
3, 3	0, 9995	0, 9995	0, 9995	0, 9996	0, 9996	0, 9996	0, 9996	0, 9996	0, 9996	0, 9997
3, 4	0, 9997	0, 9997	0, 9997	0, 9997	0, 9997	0, 9997	0, 9997	0, 9997	0, 9997	0, 9998

Tabela 2: Área de $\Phi(x)$ sob a curva normal padrão à esquerda de x.

Fonte: Arquivo do próprio autor, baseado em [15].

v	0, 50	0, 60	0, 70	0, 80	0, 90	0, 95	0, 96	0, 98	0, 99
1	1, 000	1, 376	1, 963	3, 078	6, 314	12, 706	15, 894	31, 821	63, 657
2	0, 816	1, 061	1, 386	1, 886	2, 920	4, 303	4, 849	6, 965	9, 925
3	0, 765	0, 978	1, 250	1, 638	2, 353	3, 182	3, 482	4, 541	5, 841
4	0, 741	0, 941	1, 190	1, 533	2, 132	2, 776	2, 998	3, 747	4, 604
5	0, 727	0, 920	1, 156	1, 476	2, 015	2, 571	2, 756	3, 365	4, 032
6	0, 718	0, 906	1, 134	1, 440	1, 943	2, 447	2, 612	3, 143	3, 707
7	0, 711	0, 896	1, 119	1, 415	1, 895	2, 365	2, 517	2, 998	3, 499
8	0, 706	0, 889	1, 108	1, 397	1, 860	2, 306	2, 449	2, 896	3, 355
9	0, 703	0, 883	1, 100	1, 383	1, 833	2, 262	2, 398	2, 821	3, 250
10	0, 700	0, 879	1, 093	1, 372	1, 812	2, 228	2, 359	2, 764	3, 169
11	0, 697	0, 876	1, 088	1, 363	1, 796	2, 201	2, 328	2, 718	3, 106
12	0, 695	0, 873	1, 083	1, 356	1, 782	2, 179	2, 303	2, 681	3, 055
13	0, 694	0, 870	1, 079	1, 350	1, 771	2, 160	2, 282	2, 650	3, 012
14	0, 692	0, 868	1, 076	1, 345	1, 761	2, 145	2, 264	2, 624	2, 977
15	0, 691	0, 866	1, 074	1, 341	1, 753	2, 131	2, 248	2, 602	2, 947
16	0, 690	0, 865	1, 071	1, 337	1, 746	2, 120	2, 235	2, 583	2, 921
17	0, 689	0, 863	1, 069	1, 333	1, 740	2, 110	2, 224	2, 567	2, 898
18	0, 688	0, 862	1, 067	1, 330	1, 734	2, 101	2, 214	2, 552	2, 878
19	0, 688	0, 861	1, 066	1, 328	1, 729	2, 093	2, 205	2, 539	2, 861
20	0, 687	0, 860	1, 064	1, 325	1, 725	2, 086	2, 197	2, 528	2, 845
21	0, 686	0, 859	1, 063	1, 323	1, 721	2, 080	2, 189	2, 518	2, 831
22	0, 686	0, 858	1, 061	1, 321	1, 717	2, 074	2, 183	2, 508	2, 819
23	0, 685	0, 858	1, 060	1, 319	1, 714	2, 069	2, 177	2, 500	2, 807
24	0, 685	0, 857	1, 059	1, 318	1, 711	2, 064	2, 172	2, 492	2, 797
25	0, 684	0, 856	1, 058	1, 316	1, 708	2, 060	2, 166	2, 485	2, 787
26	0, 684	0, 856	1, 058	1, 315	1, 706	2, 056	2, 162	2, 479	2, 779
27	0, 684	0, 855	1, 057	1, 314	1, 703	2, 052	2, 158	2, 473	2, 771
28	0, 684	0, 855	1, 056	1, 313	1, 701	2, 048	2, 154	2, 467	2, 763
29	0, 683	0, 854	1, 055	1, 311	1, 699	2, 045	2, 150	2, 462	2, 756
30	0, 683	0, 854	1, 055	1, 310	1, 697	2, 042	2, 147	2, 457	2, 750
35	0, 682	0, 852	1, 052	1, 306	1, 690	2, 030	2, 133	2, 438	2, 724
40	0, 681	0, 851	1, 050	1, 303	1, 684	2, 021	2, 123	2, 423	2, 704
50	0, 679	0, 849	1, 047	1, 299	1, 676	2, 009	2, 109	2, 403	2, 678
60	0, 679	0, 848	1, 045	1, 296	1, 671	2, 000	2, 099	2, 390	2, 660
120	0, 677	0, 845	1, 041	1, 289	1, 658	1, 980	2, 076	2, 358	2, 617
∞	0, 674	0, 842	1, 036	1, 282	1, 645	1, 960	2, 054	2, 326	2, 576

Tabela 3: Distribuição t de Student para um nível de confiança $1 - \alpha$.

Fonte: Arquivo do próprio autor, baseado em [5] e [8].

v	0, 99	0, 98	0, 975	0, 95	0, 90	0, 10	0, 05	0, 025	0, 02	0, 01
1	–	–	0, 001	0, 004	0, 016	2, 706	3, 841	5, 024	5, 412	6, 635
2	0, 020	0, 040	0, 051	0, 103	0, 211	4, 605	5, 991	7, 378	7, 824	9, 210
3	0, 115	0, 185	0, 216	0, 352	0, 584	6, 251	7, 815	9, 348	9, 837	11, 345
4	0, 297	0, 429	0, 484	0, 711	1, 064	7, 779	9, 488	11, 143	11, 668	13, 277
5	0, 554	0, 752	0, 831	1, 145	1, 610	9, 236	11, 070	12, 832	13, 388	15, 086
6	0, 872	1, 134	1, 237	1, 635	2, 204	10, 645	12, 592	14, 449	15, 033	16, 812
7	1, 239	1, 564	1, 690	2, 167	2, 833	12, 017	14, 067	16, 013	16, 622	18, 475
8	1, 646	2, 032	2, 180	2, 733	3, 490	13, 362	15, 507	17, 534	18, 168	20, 090
9	2, 088	2, 532	2, 700	3, 325	4, 168	14, 684	16, 919	19, 023	19, 679	21, 666
10	2, 558	3, 059	3, 247	3, 940	4, 865	15, 987	18, 307	20, 483	21, 161	23, 209
11	3, 053	3, 609	3, 816	4, 575	5, 578	17, 275	19, 675	21, 920	22, 618	24, 725
12	3, 571	4, 178	4, 404	5, 226	6, 304	18, 549	21, 026	23, 337	24, 054	26, 217
13	4, 107	4, 756	5, 009	5, 892	7, 042	19, 812	22, 362	24, 736	25, 472	27, 688
14	4, 660	5, 368	5, 629	6, 571	7, 790	21, 064	23, 685	26, 119	26, 873	29, 141
15	5, 229	5, 985	6, 262	7, 261	8, 547	22, 307	24, 996	27, 488	28, 259	30, 578
16	5, 812	6, 614	6, 908	7, 962	9, 312	23, 542	26, 296	28, 854	29, 633	32, 000
17	6, 408	7, 255	7, 564	8, 672	10, 085	24, 769	27, 587	30, 191	30, 995	33, 409
18	7, 015	7, 906	8, 231	9, 390	10, 865	25, 989	28, 869	31, 526	32, 346	34, 805
19	7, 633	8, 567	8, 906	10, 117	11, 651	27, 204	30, 144	32, 852	33, 687	36, 191
20	8, 260	9, 237	9, 591	10, 851	12, 443	28, 412	31, 410	34, 170	35, 020	37, 566
21	8, 897	9, 915	10, 283	11, 591	13, 240	29, 615	32, 671	35, 479	36, 343	38, 932
22	9, 542	10, 600	10, 982	12, 338	14, 041	30, 813	33, 924	36, 781	37, 659	40, 289
23	10, 196	11, 293	11, 688	13, 091	14, 848	32, 007	35, 172	38, 076	38, 968	41, 638
24	10, 856	11, 992	12, 401	13, 848	15, 659	33, 196	36, 415	39, 364	40, 270	42, 980
25	11, 524	12, 697	13, 120	14, 611	16, 473	34, 382	37, 652	40, 646	41, 566	44, 314
26	12, 198	13, 409	13, 844	15, 379	17, 292	35, 563	38, 885	41, 923	42, 856	45, 642
27	12, 879	14, 125	14, 573	16, 151	18, 114	36, 741	40, 113	43, 194	44, 140	46, 963
28	13, 565	14, 847	15, 308	16, 928	18, 939	37, 916	41, 337	44, 461	45, 419	48, 278
29	14, 258	15, 574	16, 047	17, 708	19, 768	39, 087	42, 557	45, 722	46, 693	49, 588
30	14, 953	16, 306	16, 971	18, 493	20, 599	40, 256	43, 773	46, 979	47, 962	50, 892

Tabela 4: Distribuição qui-quadrado, valores de α.

Fonte: Arquivo do próprio autor, baseado em [5] e [8].

Figura 45: Cães e gatos

Fonte: Arquivo do próprio autor, baseado em [19].

Crédito das imagens: http://emojipedia.org - Acessado em 30/01/16.

REFERÊNCIAS

[1] BERGER, Roger L.; CASELLA, George. *Inferência Estatística,* 2. ed. São Paulo: Cengage Learning, 2014.

[2] BOLFARINE, Heleno; SANDOVAL, Mônica C. *Introdução* à *inferência estatística,* 2. ed. Rio de Janeiro: SBM, 2010.

[3] BOYER, Carl. B. *História da Matemática,* trad. Elza F. Gomide, 3. ed. São Paulo: Blücher, 2010.

[4] BURTON, David M. *The history of mathematics : an introduction,* , 7. ed. New York: McGraw-Hill, 2011.

[5] BUSSAB, Wilton de O., MORETTIN, Pedro A. *Estatística básica,* 5. ed. São Paulo: Saraiva, 2002.

[6] CARVALHO, Paulo C. P.; MORGADO, Augusto C. *Matemática Discreta: Coleção PROFMAT,* 1. ed. Rio de Janeiro: SBM, 2014.

[7] EVES, Howard. *Introdução à história da matemática,* trad. Hygino H. Domin- gues, 5. ed. Campinas: Unicamp, 2011.

[8] FARBER, Betsy; LARSON Ron. *Estatística aplicada,* 4. ed. São Paulo: Pearson Prentice Hall, 2010.

[9] FLEMMING, Diva M.; GONÇALVES, Mírian B. *Cálculo A: funções, limite, deriva-* ção, *integração,* 5. ed. São Paulo: Makron, 2002.

[10] FREUND, John E. *Estatística aplicada: economia, adiministação e contabilidade,* 11. ed. Porto Alegre: Bookman, 2006.

[11] GRANDO, Regina C.; NACARATO, Aldair M. ; et al. *Estatística e probabilidade na educação básica: professores narrando suas experiências,* 1. ed. Campinas: Mercado de Letras, 2013.

[12] KEEPING, E. S. *Introduction to statistical inference,* New York: Dover, 1995.

[13] NACIONAIS, Parâmetros Curriculares. *Ensino Médio*. Brasília: Ministério da Educação, 2000.

[14] ROONEY, Anne. *A História da Matemática: Desde a criação das pirâmides até a exploração do infinito,*São Paulo: Mboobs, 2012.

[15] ROSS, Sheldon. *Probabilidade: um curso moderno com aplicações,* 8. ed. Porto Alegre: Bookman, 2010.

[16] SANTANA, Mario de S. *Estatística para professores da educação básica: Concei- tos e aprendizagem para a cidadania,* 1. ed. Curitiba: CRV, 2012.

[17] SCHAY, Géza. *Introduction to Probability with Statisctical Applications,* Boston: Birkäuser, 2007.

[18] STIGLER, Stephen M. *The History of Statistics: The Measurement of Uncertainty before 1900,* 9. ed. Cambridge: Harvard University Press, 2003.

[19] https://www.ime.usp.br/ativestat/atividades/aula/sa08 - Acessado em 29/01/16.

PAULO ROBERTO CONSTANTINO JUNIOR - Possui graduação em Matemática (Licenciatura) pela Universidade de Guarulhos (2002-2004), especialização em Educação Matemática pelas Faculdades Oswaldo Cruz (2009-2011), mestrado em Matemática pela Universidade Federal do ABC (2014-2016). Atua como professor de Matemática no ensino médio. ID Lattes: 2382469882819264

Ano 2023

INFERÊNCIA NO ENSINO MÉDIO:

UMA INTRODUÇÃO AOS TESTES DE HIPÓTESE

www.atenaeditora.com.br
contato@atenaeditora.com.br
@atenaeditora
www.facebook.com/atenaeditora.com.br

Ano 2023

INFERÊNCIA NO ENSINO MÉDIO:

UMA INTRODUÇÃO AOS TESTES DE HIPÓTESE

www.atenaeditora.com.br
contato@atenaeditora.com.br
@atenaeditora
www.facebook.com/atenaeditora.com.br